

PS Example 6; Page 46-47; 65pp; English.

XX AAY16615-28 represent human phosphoprotein 32 (pp32) protein sequences
CC isolated from different patients. The specification describes pp32
CC variants, designated pp32r1 and pp32r2. The pp32r1 and pp32r2 sequences
CC are associated with cancer in prostate, especially prostatic
CC adenocarcinomas. Normal pp32 exerts antineoplastic activity through
CC suppression of transformation. Cancer-associated pp32 variants augment,
CC rather than inhibit, transformation. Determining the presence of a gene
CC encoding residues 146-163 of pp32r1 or pp32r2 in a sample is useful for
CC a diagnostic method for predicting malignant potential of
CC neuroendocrine, neural, mesenchymal, lymphoid, epithelial or germ
CC cell-derived tumours.

XX SQ Sequence 234 AA;
Query Match 100.0%; Score 1216; DB 20; Length 234;
Best Local Similarity 100.0%; Pred. No. 3.5e-98;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEMGRIHSELNRAPSDVKELALDNRSGNEKLEALTDFFEFLSKINGLTSISDL 60
DB 1 MEMGRIHSELNRAPSDVKELALDNRSGNEKLEALTDFFEFLSKINGLTSISDL 60
QY 61 PKLRLKLELRVSGGLEVLAEKCPNLTLYLSGNKIKDLSLTIPLKLENKSLDLFNC 120
DB 61 PKLRLKLELRVSGGLEVLAEKCPNLTLYLSGNKIKDLSLTIPLKLENKSLDLFNC 120
QY 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDDEEGEHEEYDEDAQ 180
DB 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDDEEGEHEEYDEDAQ 180
QY 181 VVEDEEGEHEEYDEEGDDEEGYNDGEVDEDEELGEEERQKRK 234
DB 181 VVEDEEGEHEEYDEEGDDEEGYNDGEVDEDEELGEEERQKRK 234

RESULT 2

AY22800
ID AAY22800 standard; Protein; 234 AA.

XX AC AAY22800;

XX DT 26-AUG-1999 (first entry)

XX DE Amino acid sequence of phosphoprotein 32 variant pp32r1.

XX KW Phosphoprotein 32; pp32; variant; pp32r1; pp32r2; cancer; prostate;
XX KW prostatic adenocarcinoma; antineoplastic activity;
XX KW transformation suppression; malignant potential; neuroendocrine;
XX KW neural; mesenchymal; lymphoid; epithelial; germ cell; tumour.

XX OS Homo sapiens.

XX PN WO9929906-A2.

XX PD 17-JUN-1999.

XX PF 11-DEC-1998; 98WO-US26433.

XX PR 12-DEC-1997; 97US-0069677.

XX PA (UYJO) UNIV JOHNS HOPKINS.

XX PI Brody JR, Kadkol SS, Kocheavar GJ, Pasternack GR;

XX DR WPI; 1999-385626/32.

XX PT Phosphoprotein 32 (pp32) related genomic sequences

XX PS Example 2; Fig 4; 65pp; English.

XX CC The present sequence represents the phosphoprotein 32 (pp32)

CC variant pp32r1, which is isolated from human placenta. The
CC pp32r1 and pp32r2 sequences are associated with cancer in prostate,
CC especially prostatic adenocarcinomas. Normal pp32 exerts antineoplastic
CC activity through suppression of transformation. Cancer-associated pp32
CC variants augment, rather than inhibit, transformation. Determining the
CC presence of a gene encoding residues 146-163 of pp32r1 or pp32r2 in a
CC sample is useful for a diagnostic method for predicting malignant
CC potential of neuroendocrine, neural, mesenchymal, lymphoid, epithelial
CC or germ cell-derived tumours.

XX SQ Sequence 234 AA;
Query Match 100.0%; Score 1216; DB 20; Length 234;
Best Local Similarity 100.0%; Pred. No. 3.5e-98;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGNEKLEALTDFFEFLSKINGLTSISDL 60
DB 1 MEMGRIHSELNRAPSDVKELALDNRSGNEKLEALTDFFEFLSKINGLTSISDL 60
QY 61 PKLRLKLELRVSGGLEVLAEKCPNLTLYLSGNKIKDLSLTIPLKLENKSLDLFNC 120
DB 61 PKLRLKLELRVSGGLEVLAEKCPNLTLYLSGNKIKDLSLTIPLKLENKSLDLFNC 120
QY 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDDEEGEHEEYDEDAQ 180
DB 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDDEEGEHEEYDEDAQ 180
QY 181 VVEDEEGEHEEYDEEGDDEEGYNDGEVDEDEELGEEERQKRK 234
DB 181 VVEDEEGEHEEYDEEGDDEEGYNDGEVDEDEELGEEERQKRK 234

RESULT 3

AAB20649
ID AAB20649 standard; Protein; 234 AA.

XX AC AAB20649;

XX DT 15-DEC-2000 (first entry)

XX DE Human pp32 protein sequence clone KG.

XX KW Human; pp32; chromosome 15q22.3-q23; cancer; phosphoprotein; APRIL;
XX KW acidic protein rich in leucine; chromosome 15q25; tumour suppressor;
XX KW malignant; cytostatic; gene therapy.

XX OS Homo sapiens.

XX PN WO200045852-A1.

XX PD 10-AUG-2000.

XX PF 03-FEB-2000; 2000WO-US02656.

XX PR 03-FEB-1999; 99US-0118667.

XX PA (UYJO) UNIV JOHNS HOPKINS.

XX PI Pasternack GR, Bai J;

XX DR WPI; 2000-514896/46.

XX PT Treatment of cancer comprising restoration of pp32 function in
XX PT malignant cells -

XX PS Example 6; Page 49-50; 90pp; English.

XX CC The present invention describes a method (M1) for treating malignant
XX CC cells comprising restoration of pp32 function. Also described are:
XX CC (1) a method (M2) of screening to determine whether a compound is an
XX CC inducer of pp32 expression comprising measuring pp32 expression by
XX CC cells cultured in the presence and absence of the compound; and

CC (2) a method (M3) of screening to determine whether a compound is an
CC inducer of pp32 function comprising measuring protein phosphatase
CC activity in cells cultured in the presence and absence of the compound.
CC The methods are useful for treating cancer and for identifying agents
CC which may be used to treat cancer. Human pp32 is a phosphoprotein which
CC has been mapped to chromosome 15q22.3-q23. The present sequence
CC represents a human pp32 protein sequence from an example from the
CC present invention.
XX
SQ Sequence 234 AA;

Query Match 100.0%; Score 1216; DB 21; Length 234;
Best Local Similarity 100.0%; Pred. No. 3.5e-98;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEMGRRHSELNRAPSDVKELALDNRSGNEGKLEALTDFEELFSLKINGLTSIDL 60
Db 1 MEMGRRHSELNRAPSDVKELALDNRSGNEGKLEALTDFEELFSLKINGLTSIDL 60
QY 61 PKLKRKLELRVSGGLEVLAEKCPNLTLYLSGNKIKDLSTIEPLKOLENLSLDLNFCE 120
Db 61 PKLKRKLELRVSGGLEVLAEKCPNLTLYLSGNKIKDLSTIEPLKOLENLSLDLNFCE 120
QY 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDEEGEHEEYDEDAQ 180
Db 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDEEGEHEEYDEDAQ 180
QY 181 VVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVDGEDEEELGEEERGGK 234
Db 181 VVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVDGEDEEELGEEERGGK 234

RESULT 4
AAB20655
ID AAB20655 standard; Protein; 234 AA.

XX AAB20655;
XX
DT 15-DEC-2000 (first entry)
XX
DE Human variant pp32r1 protein sequence.

XX Human; pp32; chromosome 15q22.3-q23; cancer; phosphoprotein; APRIL;
KW acidic protein rich in leucine; Chromosome 15q25; tumour suppressor;
KW malignant; cytostatic; gene therapy.
XX
OS Homo sapiens.

XX WO200045852-A1.
XX
XX 10-AUG-2000.
XX
XX 03-FEB-2000; 2000WO-US02656.
XX
XX 03-FEB-1999; 99US-0118667.
XX
XX (UYJO) UNIV JOHNS HOPKINS.
XX
XX Pasternack GR, Bai J;

XX WPI; 2000-514896/46.
XX N-PSDB; AAA88237.
XX
XX Treatment of cancer comprising restoration of pp32 function in
XX malignant cells -
XX
XX Example 3; Fig 4; 90pp; English.

XX The present invention describes a method (M1) for treating malignant
XX cells comprising restoration of pp32 function. Also described are:
CC (1) a method (M2) of screening to determine whether a compound is an
CC inducer of pp32 expression comprising measuring pp32 expression by
CC cells cultured in the presence and absence of the compound; and

CC (2) a method (M3) of screening to determine whether a compound is an
CC inducer of pp32 function comprising measuring protein phosphatase
CC activity in cells cultured in the presence and absence of the compound.
CC The methods are useful for treating cancer and for identifying agents
CC which may be used to treat cancer. Human pp32 is a phosphoprotein which
CC has been mapped to chromosome 15q22.3-q23. The present sequence
CC represents the human variant pp32r1 protein sequence from
CC an example of the present invention.
XX
SQ Sequence 234 AA;

Query Match 100.0%; Score 1216; DB 21; Length 234;
Best Local Similarity 100.0%; Pred. No. 3.5e-98;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEMGRRHSELNRAPSDVKELALDNRSGNEGKLEALTDFEELFSLKINGLTSIDL 60
Db 1 MEMGRRHSELNRAPSDVKELALDNRSGNEGKLEALTDFEELFSLKINGLTSIDL 60
QY 61 PKLKRKLELRVSGGLEVLAEKCPNLTLYLSGNKIKDLSTIEPLKOLENLSLDLNFCE 120
Db 61 PKLKRKLELRVSGGLEVLAEKCPNLTLYLSGNKIKDLSTIEPLKOLENLSLDLNFCE 120
QY 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDEEGEHEEYDEDAQ 180
Db 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDEEGEHEEYDEDAQ 180
QY 181 VVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVDGEDEEELGEEERGGK 234
Db 181 VVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVDGEDEEELGEEERGGK 234

RESULT 5
AAV16622
ID AAV16622 standard; Protein; 234 AA.

XX AAV16622;
XX
DT 26-AUG-1999 (first entry)
XX
DE Human phosphoprotein 32 (pp32) protein sequence.

XX Phosphoprotein 32; pp32; variant; pp32r1; pp32r2; cancer; prostate;
KW prostatic adenocarcinoma; antineoplastic activity; PCR primer;
KW transformation suppression; malignant potential; neuroendocrine;
KW neural; mesenchymal; lymphoid; epithelial; germ cell; tumour.

OS Homo sapiens.
XX
XX WO929906-A2.
XX
XX 17-JUN-1999.
XX
XX 11-DEC-1998; 98WO-US26433.
XX
XX 12-DEC-1997; 97US-0069677.
XX
XX (UYJO) UNIV JOHNS HOPKINS.
XX
XX Brody JR, Kadkol SS, Kochevar GJ, Pasternack GR;
XX WPI; 1999-385626/32.
XX
XX Phosphoprotein 32 (pp32) related genomic sequences
XX
XX Example 6; Page 46-47; 65pp; English.

XX AAV16615-28 represent human phosphoprotein 32 (pp32) protein sequences
XX isolated from different patients. The specification describes pp32
XX variants, designated pp32r1 and pp32r2. The pp32r1 and pp32r2 sequences
XX are associated with cancer in prostate, especially prostatic
XX adenocarcinomas. Normal pp32 exerts antineoplastic activity through
XX suppression of transformation. Cancer-associated pp32 variants augment,

CC rather than inhibit, transformation. Determining the presence of a gene
CC encoding residues 146-163 of pp32r1 or pp32r2 in a sample is useful for
CC a diagnostic method for predicting malignant potential of
CC neuroendocrine, neural, mesenchymal, lymphoid, epithelial or germ
CC cell-derived tumours.

XX SQ Sequence 234 AA;
Query Match 99.4%; Score 1209; DB 20; Length 234;
Best Local Similarity 99.1%; Pred. No. 1.4e-97;
Matches 232; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEMGRIHSELNRAPSDVKELALDNRSGNECKLEALTDDEFELEFLSKINGLTSIDL 60
DB 1 MEMGRIHSELNRAPSDVKELVLDNRSGNECKLEALTDDEFELEFLSKINGLTSIDL 60
QY 61 PKLKLRLKLELVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKOLENLSLDLNFCE 120
DB 61 PKLKLRLKLELVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKOLENLSLDLNFCE 120
QY 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSIDEDHVEGLDDEEGEHEEYDEDAQ 180
DB 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSIDEDHVEGLDDEEGEHEEYDEDAQ 180
QY 181 VVEDEGEHEEYDEGDEDEGYNDGVDGEDEEELGEEERQKRR 234
DB 181 VVEDEGEHEEYDEGDEDEGYNDGVDGEDEEELGEEERQKRR 234

RESULT 6
AAB20648
ID AAB20648 standard; Protein; 234 AA.
XX AC AAB20648;
XX DT 15-DEC-2000 (first entry)
XX DE Human pp32 protein sequence clone FT2.2.
XX KW Human; pp32; chromosome 15q22.3-q23; cancer; phosphoprotein; APRIL;
KW acidic protein rich in leucine; chromosome 15q25; tumour suppressor;
KW malignant; cytostatic; gene therapy.
XX OS Homo sapiens.
XX PN WO200045852-A1.
XX PD 10-AUG-2000.
XX PF 03-FEB-2000; 2000WO-US02656.
XX PR 03-FEB-1999; 99US-0118667.
XX PA (UYJO) UNIV JOHNS HOPKINS.
XX PI Pasternack GR, Bai J;
XX DR WPI; 2000-514896/46.
XX PT Treatment of cancer comprising restoration of pp32 function in
XX malignant cells -
XX PS Example 6; Page 49-50; 90pp; English.
XX CC The present invention describes a method (M1) for treating malignant
CC cells comprising restoration of pp32 function. Also described are:
CC (1) a method (M2) of screening to determine whether a compound is an
CC inducer of pp32 expression comprising measuring pp32 expression by
CC cells cultured in the presence and absence of the compound; and
CC (2) a method (M3) of screening to determine whether a compound is an
CC inducer of pp32 function comprising measuring protein phosphatase
CC activity in cells cultured in the presence and absence of the compound.
CC The methods are useful for treating cancer and for identifying agents

CC which may be used to treat cancer. Human pp32 is a phosphoprotein which
CC has been mapped to chromosome 15q22.3-q23. The present sequence
CC represents a human pp32 protein sequence from an example from the
CC present invention.

XX SQ Sequence 234 AA;
Query Match 99.4%; Score 1209; DB 21; Length 234;
Best Local Similarity 99.1%; Pred. No. 1.4e-97;
Matches 232; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEMGRIHSELNRAPSDVKELALDNRSGNECKLEALTDDEFELEFLSKINGLTSIDL 60
DB 1 MEMGRIHSELNRAPSDVKELVLDNRSGNECKLEALTDDEFELEFLSKINGLTSIDL 60
QY 61 PKLKLRLKLELVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKOLENLSLDLNFCE 120
DB 61 PKLKLRLKLELVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKOLENLSLDLNFCE 120
QY 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSIDEDHVEGLDDEEGEHEEYDEDAQ 180
DB 121 VTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSIDEDHVEGLDDEEGEHEEYDEDAQ 180
QY 181 VVEDEGEHEEYDEGDEDEGYNDGVDGEDEEELGEEERQKRR 234
DB 181 VVEDEGEHEEYDEGDEDEGYNDGVDGEDEEELGEEERQKRR 234

RESULT 7
AA116624
ID AAX16624 standard; Protein; 245 AA.
XX AC AAX16624;
XX DT 26-AUG-1999 (first entry)
XX DE Human phosphoprotein 32 (pp32) protein sequence.
XX KW Phosphoprotein 32; pp32; variant; pp32r1; pp32r2; cancer; prostate;
KW prostatic adenocarcinoma; antineoplastic activity; PCR primer;
KW transformation suppression; malignant potential; neuroendocrine;
KW neural; mesenchymal; lymphoid; epithelial; germ cell; tumour.
XX OS Homo sapiens.
XX PN WO9929906-A2.
XX PD 17-JUN-1999.
XX PF 11-DEC-1998; 98WO-US26433.
XX PR 12-DEC-1997; 97US-0069677.
XX PA (UYJO) UNIV JOHNS HOPKINS.
XX PI Brody JR, Kadkol SS, Kocheavar GJ, Pasternack GR;
XX DR WPI; 1999-385626/32.
XX PT Phosphoprotein 32 (pp32) related genomic sequences
XX PS Example 6; Page 46-47; 65pp; English.
XX CC AAX16615-28 represent human phosphoprotein 32 (pp32) protein sequences
CC isolated from different patients. The specification describes pp32
CC variants, designated pp32r1 and pp32r2. The pp32r1 and pp32r2 sequences
CC are associated with cancer in prostate, especially prostatic
CC adenocarcinomas. Normal pp32 exerts antineoplastic activity through
CC suppression of transformation. Cancer-associated pp32 variants augment,
CC rather than inhibit, transformation. Determining the presence of a gene
CC encoding residues 146-163 of pp32r1 or pp32r2 in a sample is useful for
CC a diagnostic method for predicting malignant potential of
CC neuroendocrine, neural, mesenchymal, lymphoid, epithelial or germ

CC cell-derived tumours.

```
XX SQ Sequence 245 AA;
Query Match 95.5%; Score 1161; DB 20; Length 245;
Best Local Similarity 95.7%; Pred. No. 2.4e-93;
Matches 224; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
QY 1 MEMGRRTHSELNRAPSDVKELALDNRSGNEKLEALTDPEFELEPLSKINGLTSIDL 60
Db 1 MEMGRRHLELRNRTPSDKVELVDNRSGNEKLEGLTDEFEFELEFLSTINVGLTSIANL 60
QY 61 PKLKRKLELRVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKQLENKSLDLFNCE 120
Db 61 PKLKRKLELRVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKQLENKSLDLFNCE 120
QY 121 VTNLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDIEDHVEGLDDEEGEHEEYDEDAQ 180
Db 121 VTNLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDIEDHVEGLDDEEGEHEEYDEDAQ 180
QY 181 VVEDEGEGEREEDVSGGDEDEEGYNDGEVDEDEELGEEERGGQRK 234
Db 181 VVEDEGEGEREEDVSGGDEDEEGYNDGEVDEDEELGEEERGGQRK 234
```

RESULT 8

```
AAB20650
ID AAB20650 standard; Protein; 245 AA.
```

AC AAB20650;

DT 15-DEC-2000 (first entry)

DE Human pp32 protein sequence clone FT1.7.

```
XX Human; pp32; chromosome 15q22.3-q23; cancer; phosphoprotein; APRIL;
KW acidic protein rich in leucine; chromosome 15q25; tumour suppressor;
KW malignant; cytostatic; gene therapy.
```

OS Homo sapiens.

PN W0200045852-A1.

PD 10-AUG-2000.

PF 03-FEB-2000; 2000WO-US02656.

PR 03-FEB-1999; 99US-0118667.

XX (UYJO) UNIV JOHNS HOPKINS.

PA Pasternack GR, Bai J;

DR WPI; 2000-514896/46.

```
XX Treatment of cancer comprising restoration of pp32 function in
PT malignant cells -
```

XX Example 6; Page 49-50; 90pp; English.

```
XX The present invention describes a method (M1) for treating malignant
CC cells comprising restoration of pp32 function. Also described are:
CC (1) a method (M2) of screening to determine whether a compound is an
CC inducer of pp32 expression comprising measuring pp32 expression by
CC cells cultured in the presence and absence of the compound; and
CC (2) a method (M3) of screening to determine whether a compound is an
CC inducer of pp32 function comprising measuring protein phosphatase
CC activity in cells cultured in the presence and absence of the compound.
CC The methods are useful for treating cancer and for identifying agents
CC which may be used to treat cancer. Human pp32 is a phosphoprotein which
CC has been mapped to chromosome 15q22.3-q23. The present sequence
CC represents a human pp32 protein sequence from an example from the
CC present invention.
```

```
XX SQ Sequence 245 AA;
Query Match 95.5%; Score 1161; DB 21; Length 245;
Best Local Similarity 95.7%; Pred. No. 2.4e-93;
Matches 224; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 MEMGRRTHSELNRAPSDVKELALDNRSGNEKLEALTDPEFELEPLSKINGLTSIDL 60
Db 1 MEMGRRHLELRNRTPSDKVELVDNRSGNEKLEGLTDEFEFELEFLSTINVGLTSIANL 60
QY 61 PKLKRKLELRVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKQLENKSLDLFNCE 120
Db 61 PKLKRKLELRVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKQLENKSLDLFNCE 120
QY 121 VTNLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDIEDHVEGLDDEEGEHEEYDEDAQ 180
Db 121 VTNLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDIEDHVEGLDDEEGEHEEYDEDAQ 180
QY 181 VVEDEGEGEREEDVSGGDEDEEGYNDGEVDEDEELGEEERGGQRK 234
Db 181 VVEDEGEGEREEDVSGGDEDEEGYNDGEVDEDEELGEEERGGQRK 234
```

RESULT 9

```
AAR95900
ID AAR95900 standard; Protein; 249 AA.
```

AC AAR95900;

DT 30-JUL-1996. (first entry)

DE Human pp32 protein.

```
XX pp32; cancer; diagnosis; therapy; antisense; cell proliferation;
KW lymphoid tumour; epithelial tumour; colon carcinoma;
KW prostate carcinoma; non-Hodgkin lymphoma; apoptosis.
```

OS Homo sapiens.

```
XX Key Location/Qualifiers
FH Domain 1..167
```

```
FT /note= "N-terminal domain is generally amphipathic
FT and has a high probability of alpha-helix
FT conformation"
```

FT Region 69..90

```
FT /note= "candidate leucine zipper composed of
FT Leu-69, Val-76, Leu-83 and Leu-90"
```

FT Region 60..68

```
FT /note= "candidate nuclear localisation sequence"
FT 108..116
```

```
FT /note= "candidate nuclear localisation sequence"
FT 168..249
```

FT /note= "highly acidic C-terminal domain"

XX W09610092-A1.

XX 04-APR-1996.

XX 28-SEP-1995; 95WO-US12414.

XX 28-SEP-1994; 94US-0314503.

XX (UYJO) UNIV JOHNS HOPKINS.

XX KuhaJda FP, Pasternack GR;

XX WPI; 1996-200930/20.

XX N-PSDB; AAT27712.

```
XX New method of diagnosing cancer using pp32 cDNA - by detecting the
PT level of mRNA hybridising to pp32 cDNA; also for inhibiting cell
PT proliferation and screening anti-cancer drugs.
```

XX Example 12; Fig 10a; 129pp; English.
 PS A 32 kDa human protein, pp32 (AAR95900), and immunologically
 CC related proteins pp35 and pp42 are useful in the diagnosis and
 CC prognosis of tumours of lymphoid and epithelial origin. pp32 is
 CC overexpressed in many neoplastic cells. Its activities include:
 CC inhibition of co-transformation by ras and myc oncogenes; partial
 CC protection against programmed cell death; and modulation of
 CC nuclear shape and size. Its sequence was deduced from a cDNA clone
 CC (AAT27712) isolated from human HL-60 cells. Transfected mammalian
 CC cells that express pp32 may be used to screen for anti-cancer drugs.
 CC Purified pp32 can be used to raise diagnostic antibodies.
 XX Sequence 249 AA;
 SQ Query Match 82.3%; Score 1001; DB 17; Length 249;
 Best Local Similarity 86.1%; Pred. No. 2.2e-79;
 Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;
 QY 1 MEMGRRTHSELNRAPSDVKELALDNRSGNEKLEALTDFFEEFLSKINGLTSIDL 60
 Db 1 MEMGRRTHLELRNRTPSDKELVDNRSGNEKLEGLTDFEEFLSKINGLTSIANL 60
 QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSSTIEPLKLENKLSIDL 116
 Db 61 PKLNLKLELSDNRVSGGLEVLAEKCPNLTHLYLSGNKIKDLSSTIEPLKLENKLSIDL 120
 QY 117 FNCVETNLNDYGENVFKLLQLTYLDSYWDHKEAPYSDIEDHVEGLDDEEHEEYD 176
 Db 61 FNCVETNLNDYGENVFKLLQLTYLDSYWDHKEAPYSDIEDHVEGLDDEEHEEYD 180
 QY 177 EDAQVVEDEGEDEEEDVSGGDEDEEGYNDGEVDGEDEELGEEERGQKRK 234
 Db 181 EDAQVVEDEGEDEEEDVSGGDEDEEGYNDGEVDGEDEELGEEERGQKRK 238

RESULT 10
 AAY16626
 ID AAY16626 standard; Protein; 249 AA.
 XX AC AAY16626;
 XX DT 26-AUG-1999 (first entry)
 XX DE Human phosphoprotein 32 (pp32) protein sequence.
 XX KW Phosphoprotein 32; pp32; variant; pp32r1; pp32r2; cancer; prostate;
 KW prostatic adenocarcinoma; antineoplastic activity; PCR primer;
 KW transformation suppression; malignant potential; neuroendocrine;
 KW neural; mesenchymal; lymphoid; epithelial; germ cell; tumour.
 XX OS Homo sapiens.
 XX PN WO9929906-A2.
 XX PD 17-JUN-1999.
 XX PF 11-DEC-1998; 98WO-US26433.
 XX PR 12-DEC-1997; 97US-0069677.
 XX PA (UYJO) UNIV JOHNS HOPKINS.
 XX PI Brody JR, Kadkol SS, Kocheavar GJ, Pasternack GR;
 XX DR WPI; 1999-385626/32.
 XX PT Phosphoprotein 32 (pp32) related genomic sequences
 XX PS Example 6; Page 46-47; 65pp; English.
 XX CC AAY16615-28 represent human phosphoprotein 32 (pp32) protein sequences

CC isolated from different patients. The specification describes pp32
 CC variants, designated pp32r1 and pp32r2. The pp32r1 and pp32r2 sequences
 CC are associated with cancer in prostate, especially prostatic
 CC adenocarcinomas. Normal pp32 exerts antineoplastic activity through
 CC suppression of transformation. Cancer-associated pp32 variants augment,
 CC rather than inhibit, transformation. Determining the presence of a gene
 CC encoding residues 146-163 of pp32r1 or pp32r2 in a sample is useful for
 CC a diagnostic method for predicting malignant potential of
 CC neuroendocrine, neural, mesenchymal, lymphoid, epithelial or germ
 CC cell-derived tumours.

XX Sequence 249 AA;
 SQ Query Match 82.3%; Score 1001; DB 20; Length 249;
 Best Local Similarity 86.1%; Pred. No. 2.2e-79;
 Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;
 QY 1 MEMGRRTHSELNRAPSDVKELALDNRSGNEKLEALTDFFEEFLSKINGLTSIDL 60
 Db 1 MEMGRRTHLELRNRTPSDKELVDNRSGNEKLEGLTDFEEFLSKINGLTSIANL 60
 QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSSTIEPLKLENKLSIDL 116
 Db 61 PKLNLKLELSDNRVSGGLEVLAEKCPNLTHLYLSGNKIKDLSSTIEPLKLENKLSIDL 120
 QY 117 FNCVETNLNDYGENVFKLLQLTYLDSYWDHKEAPYSDIEDHVEGLDDEEHEEYD 176
 Db 121 FNCVETNLNDYGENVFKLLQLTYLDSYWDHKEAPYSDIEDHVEGLDDEEHEEYD 180
 QY 177 EDAQVVEDEGEDEEEDVSGGDEDEEGYNDGEVDGEDEELGEEERGQKRK 234
 Db 181 EDAQVVEDEGEDEEEDVSGGDEDEEGYNDGEVDGEDEELGEEERGQKRK 238

RESULT 11
 AAY16627
 ID AAY16627 standard; Protein; 249 AA.
 XX AC AAY16627;
 XX DT 26-AUG-1999 (first entry)
 XX DE Human phosphoprotein 32 (pp32) protein sequence.
 XX KW Phosphoprotein 32; pp32; variant; pp32r1; pp32r2; cancer; prostate;
 KW prostatic adenocarcinoma; antineoplastic activity; PCR primer;
 KW transformation suppression; malignant potential; neuroendocrine;
 KW neural; mesenchymal; lymphoid; epithelial; germ cell; tumour.
 XX OS Homo sapiens.
 XX PN WO9929906-A2.
 XX PD 17-JUN-1999.
 XX PF 11-DEC-1998; 98WO-US26433.
 XX PR 12-DEC-1997; 97US-0069677.
 XX PA (UYJO) UNIV JOHNS HOPKINS.
 XX PI Brody JR, Kadkol SS, Kocheavar GJ, Pasternack GR;
 XX DR WPI; 1999-385626/32.
 XX PT Phosphoprotein 32 (pp32) related genomic sequences
 XX PS Example 6; Page 46-47; 65pp; English.
 XX CC AAY16615-28 represent human phosphoprotein 32 (pp32) protein sequences
 CC isolated from different patients. The specification describes pp32
 CC variants, designated pp32r1 and pp32r2. The pp32r1 and pp32r2 sequences
 CC are associated with cancer in prostate, especially prostatic

PT New isolated V3 loop HIV receptor - comprises P95/nucleolin, for the
PT P40/PHAPII and P30/PHAPI proteins, used to develop products for the
XX treatment and prevention of HIV infection
XX
XX Disclosure; Fig 49(11); 267pp; English.

XX This is the amino acid sequence of the P30 (PHAPI) protein of
CC the newly identified V3 loop HIV receptor. This novel protein
CC complex receptor for HIV retroviruses consists of an association of
CC 3 proteins, named P95/nucleolin, P40/PHAPII and P30/PHAPI (see
CC AAW84052-54); that are implicated as cofactors in the process of
CC HIV entry into cells. Genomic and cDNA sequences for these proteins
CC are provided in AAW71742-45. The V3 loop HIV receptor proteins were
CC isolated from human CD4+ CEM T-cell extracts using an affinity
CC matrix containing either the pseudopeptide 5(KPSICH2)NPR-template
CC assembled synthetic peptide or a synthetic V3 loop peptide (see
CC AAW84055). P30 was PHAPI as nucleolin by amino acid sequence
CC analysis. The invention also concerns peptidic or non-peptidic
CC molecules having the ability to alter and/or prevent the binding of
CC the novel HIV receptor to the HIV retrovirus, and to pharmaceutical
CC and diagnostic compositions containing such molecules. Methods are
CC provided for screening for new active molecules, and to methods of
CC screening genetic defects in the expression of the V3 loop HIV
CC receptor in individuals that survive long-term HIV infection or who
CC are HIV-resistant. Such genetically defective polynucleotides can
CC be used in gene therapy.

XX Sequence 249 AA;

Query Match 82.3%; Score 1001; DB 20; Length 249;
Best Local Similarity 86.1%; Pred. No. 2.2e-79;
Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;

QY 1 MEMGRRTHSELNRAPSDVKELALDNRSGKLEALTDDEFEFLSKINGLTSIDL 60
DB 1 MEMGRRTHSELNRAPSDVKELALDNRSGKLEALTDDEFEFLSKINGLTSIDL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLGNKIKDLSTIEPLKLENLSLDL 116
DB 61 PKLNLKLELSDNRVSGGLEVLAEKCPNLTLYLGNKIKDLSTIEPLKLENLSLDL 120
QY 117 FNCVNTLNNDYGENVFKLLQLTYLDCYWDHKEAPYSDIEDHVGDLDEEGHEEYD 176
DB 121 FNCVNTLNNDYGENVFKLLQLTYLDCYWDHKEAPYSDIEDHVGDLDEEGHEEYD 180
QY 177 EDAQVVEDEEEDVEEGEEDVSGDEEEDYNDGEVDEEDELGEERGGQRK 234
DB 181 EDAQVVEDEEEDVEEGEEDVSGDEEEDYNDGEVDEEDELGEERGGQRK 238

RESULT 14
AAB20652
ID AAB20652 standard; Protein; 249 AA.

XX AAB20652;
XX
XX 15-DEC-2000 (first entry)
XX
XX Human pp32 protein sequence clone L3.

XX Human; pp32; chromosome 15q22.3-q23; cancer; phosphoprotein; APRIL;
KW acidic protein rich in leucine; chromosome 15q25; tumour suppressor;
KW malignant; cytostatic; gene therapy.

XX Homo sapiens.

XX WO200045852-A1.

XX 10-AUG-2000.

XX 03-FEB-2000; 2000WO-US02656.

XX 03-FEB-1999; 99US-0118667.

XX (UYJO) UNIV JOHNS HOPKINS.
PA Pasternack GR, Bai J;
XX
XX WPI; 2000-514896/46.

XX Treatment of cancer comprising restoration of pp32 function in
PT malignant cells -
XX
XX Example 6; Page 49-50; 90pp; English.

XX The present invention describes a method (M1) for treating malignant
CC cells comprising restoration of pp32 function. Also described are:
CC (1) a method (M2) of screening to determine whether a compound is an
CC inducer of pp32 expression comprising measuring pp32 expression by
CC cells cultured in the presence and absence of the compound; and
CC (2) a method (M3) of screening to determine whether a compound is an
CC inducer of pp32 function comprising measuring protein phosphatase
CC activity in cells cultured in the presence and absence of the compound.
CC The methods are useful for treating cancer and for identifying agents
CC which may be used to treat cancer. Human pp32 is a phosphoprotein which
CC has been mapped to chromosome 15q22.3-q23. The present sequence
CC represents a human pp32 protein sequence from an example from the
CC present invention.

XX Sequence 249 AA;

Query Match 82.3%; Score 1001; DB 21; Length 249;
Best Local Similarity 86.1%; Pred. No. 2.2e-79;
Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;

QY 1 MEMGRRTHSELNRAPSDVKELALDNRSGKLEALTDDEFEFLSKINGLTSIDL 60
DB 1 MEMGRRTHSELNRAPSDVKELALDNRSGKLEALTDDEFEFLSKINGLTSIDL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLGNKIKDLSTIEPLKLENLSLDL 116
DB 61 PKLNLKLELSDNRVSGGLEVLAEKCPNLTLYLGNKIKDLSTIEPLKLENLSLDL 120
QY 117 FNCVNTLNNDYGENVFKLLQLTYLDCYWDHKEAPYSDIEDHVGDLDEEGHEEYD 176
DB 121 FNCVNTLNNDYGENVFKLLQLTYLDCYWDHKEAPYSDIEDHVGDLDEEGHEEYD 180
QY 177 EDAQVVEDEEEDVEEGEEDVSGDEEEDYNDGEVDEEDELGEERGGQRK 234
DB 181 EDAQVVEDEEEDVEEGEEDVSGDEEEDYNDGEVDEEDELGEERGGQRK 238

RESULT 15
AAB20653
ID AAB20653 standard; Protein; 249 AA.

XX AAB20653;
XX
XX 15-DEC-2000 (first entry)
XX
XX Human pp32 protein sequence clone pp32.

XX Human; pp32; chromosome 15q22.3-q23; cancer; phosphoprotein; APRIL;
KW acidic protein rich in leucine; chromosome 15q25; tumour suppressor;
KW malignant; cytostatic; gene therapy.

XX Homo sapiens.

XX WO200045852-A1.

XX 10-AUG-2000.

XX 03-FEB-2000; 2000WO-US02656.

XX 03-FEB-1999; 99US-0118667.


```
Db 1 MEMGRIHLELRNRTSDVKELVLDNSRNEKLEGLTDFEFLSTINVGLTSIDL 60
Qy 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKOLENKSIDL 116
Db 61 PKLKLKLELSDNRVSGGLEVLAEKCPNLTHLNSGNKIKDLSTIEPLKLENKSIDL 120
Qy 117 FNCVTNLNDYGNVFKLLQLTYLSDCYWDHKEAPYSIDHVEGLDDEEGEHEEYD 176
Db 121 FNCVTNLNDYGNVFKLLPQTYLDGYDRDKEAPDSAEGYVEGLDDEEDED 180
Qy 177 EDAQVVEDEEGEHEEEDVSGGDEDEEGYNDGVDGDEDEEELGEEERQKRK 234
Db 181 EDAQVVEDEDEEGEHEEEDVSGGDEDEEGYNDGVDGDEDEEELGEEERQKRK 238

RESULT 2
US-08-314-503A-2
; Sequence 2, Application US/08314503A
; Patent No. 5734022
; GENERAL INFORMATION:
; APPLICANT: Pasternack, Gary R.
; APPLICANT: Kuhaajda, Francis P.
; TITLE OF INVENTION: No. 5734022el Mammalian Protein Associated With
; TITLE OF INVENTION: Uncontrolled Cell Division
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/314,503A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske Esq., Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 1107.47218
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9153
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 249 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-314-503A-2

Query Match 82.3%; Score 1001; DB 1; Length 249;
Best Local Similarity 86.1%; Pred. No. 3.5e-88;
Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;

Qy 1 MEMGRIHSELRNAPSVDKELALDNSRNEKLEALTDFEFLSKINGLTSIDL 60
Db 1 MEMGRIHLELRNRTSDVKELVLDNSRNEKLEGLTDFEFLSTINVGLTSIDL 60
Qy 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKOLENKSIDL 116
Db 61 PKLKLKLELSDNRVSGGLEVLAEKCPNLTHLNSGNKIKDLSTIEPLKLENKSIDL 120
Qy 117 FNCVTNLNDYGNVFKLLQLTYLSDCYWDHKEAPYSIDHVEGLDDEEGEHEEYD 176
Db 121 FNCVTNLNDYGNVFKLLPQTYLDGYDRDKEAPDSAEGYVEGLDDEEDED 180
Qy 177 EDAQVVEDEEGEHEEEDVSGGDEDEEGYNDGVDGDEDEEELGEEERQKRK 234
Db 181 EDAQVVEDEDEEGEHEEEDVSGGDEDEEGYNDGVDGDEDEEELGEEERQKRK 238

RESULT 4
US-08-466-717-2
; Sequence 2, Application US/08466717
; Patent No. 5874234
```

```
Db 181 EDAQVVEDEDEEGEHEEEDVSGGDEDEEGYNDGVDGDEDEEELGEEERQKRK 238

RESULT 3
US-08-468-066-2
; Sequence 2, Application US/08468066
; Patent No. 5756676
; GENERAL INFORMATION:
; APPLICANT: Pasternack, Gary R.
; APPLICANT: Kuhaajda, Francis P.
; TITLE OF INVENTION: No. 5756676el Mammalian Protein Associated With
; TITLE OF INVENTION: Uncontrolled Cell Division
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,066
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,503
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske Esq., Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 1107.47218
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9153
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 249 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-066-2

Query Match 82.3%; Score 1001; DB 1; Length 249;
Best Local Similarity 86.1%; Pred. No. 3.5e-88;
Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;

Qy 1 MEMGRIHSELRNAPSVDKELALDNSRNEKLEALTDFEFLSKINGLTSIDL 60
Db 1 MEMGRIHLELRNRTSDVKELVLDNSRNEKLEGLTDFEFLSTINVGLTSIDL 60
Qy 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKOLENKSIDL 116
Db 61 PKLKLKLELSDNRVSGGLEVLAEKCPNLTHLNSGNKIKDLSTIEPLKLENKSIDL 120
Qy 117 FNCVTNLNDYGNVFKLLQLTYLSDCYWDHKEAPYSIDHVEGLDDEEGEHEEYD 176
Db 121 FNCVTNLNDYGNVFKLLPQTYLDGYDRDKEAPDSAEGYVEGLDDEEDED 180
Qy 177 EDAQVVEDEEGEHEEEDVSGGDEDEEGYNDGVDGDEDEEELGEEERQKRK 234
Db 181 EDAQVVEDEDEEGEHEEEDVSGGDEDEEGYNDGVDGDEDEEELGEEERQKRK 238

RESULT 4
US-08-466-717-2
; Sequence 2, Application US/08466717
; Patent No. 5874234
```


COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,743
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/314,503
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Posorske Esq., Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 1107.47218
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-743-2

Query Match 82.3%; Score 1001; DB 3; Length 249;
Best Local Similarity 86.1%; Pred. No. 3.5e-88;
Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;
QY 1 MEMGRIHSELNRAPSDVKELALDNRSGKLEALTDFFELEFLSKINGLTSIDL 60
DB 1 MEMGRIHLELRNTPSDVKELVLDNRSGKLEGLTDFEELFSLTINVGLTSIANL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLHSGNKIKDLSLTPLEKLENKSLDL 116
DB 61 PKLNLKLELSDNRVSGGLEVLAEKCPNLTLYLHSGNKIKDLSLTPLEKLENKSLDL 120
QY 117 FNCVNTLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDEDHVGSLDDEEGEHEEYD 176
DB 121 FNCVNTLNDYRENVFKLLPQLTYLDYDRDKAPSDAEGYVGLDDDEDEDEEYD 180
QY 177 EDQVVEDEGEDEEEDVSGDEDEEGYNDGEVDGDEDEELGEEERGGK 234
DB 181 EDQVVEDEDEDEEEDVSGDEDEEGYNDGEVDGDEDEELGEEERGGK 238

RESULT 7
US-09-262-610-4
Sequence 4, Application US/09262610
Patent No. 642849
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/262,610
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/766,738
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0177 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 403007
US-09-262-610-4

Query Match 82.3%; Score 1001; DB 4; Length 249;
Best Local Similarity 86.1%; Pred. No. 3.5e-88;
Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;
QY 1 MEMGRIHSELNRAPSDVKELALDNRSGKLEALTDFFELEFLSKINGLTSIDL 60
DB 1 MEMGRIHLELRNTPSDVKELVLDNRSGKLEGLTDFEELFSLTINVGLTSIANL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLHSGNKIKDLSLTPLEKLENKSLDL 116
DB 61 PKLNLKLELSDNRVSGGLEVLAEKCPNLTLYLHSGNKIKDLSLTPLEKLENKSLDL 120
QY 117 FNCVNTLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDEDHVGSLDDEEGEHEEYD 176
DB 121 FNCVNTLNDYRENVFKLLPQLTYLDYDRDKAPSDAEGYVGLDDDEDEDEEYD 180
QY 177 EDQVVEDEGEDEEEDVSGDEDEEGYNDGEVDGDEDEELGEEERGGK 234
DB 181 EDQVVEDEDEDEEEDVSGDEDEEGYNDGEVDGDEDEELGEEERGGK 238

RESULT 8
PCT-US95-12414-2
Sequence 2, Application PC/TUS9512414
GENERAL INFORMATION:
APPLICANT: Pasternack, Gary R.
APPLICANT: Kuhajda, Francis P.
TITLE OF INVENTION: Novel Mammalian Protein Associated With
TITLE OF INVENTION: Uncontrolled Cell Division
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12414
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,503
FILING DATE: 22-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hoscheit Esq., Dale H.

REGISTRATION NUMBER: 19,090
REFERENCE/DOCKET NUMBER: 1107.51507
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-12414-2

Query Match 82.3%; Score 1001; DB 5; Length 249;

Best Local Similarity 86.1%; Pred. No. 3.5e-98;
Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;

QY 1 MEMGRRTHSELNRAPSDVKELALDNRSGKLEALTDFEEFLSKINGLTSISDL 60
DB 1 MEMGRRTHSELNRTPSDVKELVLDNRSGKLEGLTDFEEFLSTINVLTSIANL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSLTIPLKQLENKSLDL 116
DB 61 PKLKLKLELSNVRSGGLEVLAEKCPNLTHLNLGNKIKDLSLTIPLKLENKSLDL 120
QY 117 FNCVNTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDEEGEHEEYD 176
DB 121 FNCVNTNLNDYRENVFKLLPQLTYLDGYDDKAPSDAEAGYVGLDDEDEDEEYD 180
QY 177 EDAQVVEDEGEDEEVEEEDVSGGDEDEEGYNDGEVDGEDEEELGEERGGKRR 234
DB 181 EDAQVVEDEDEDEEVEEEDVSGEEDVEGYNDGEVDDEDEEELGEERGGKRR 238

RESULT 9

US-08-766-738-3
; Sequence 3, Application US/08766738
; Patent No. 5916749
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,738
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0177 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1498225
US-08-766-738-3

Query Match 56.2%; Score 683.5; DB 2; Length 251;
Best Local Similarity 61.0%; Pred. No. 8.7e-58;
Matches 152; Conservative 34; Mismatches 40; Indels 23; Gaps 8;

QY 1 MEMGRRTHSELNRAPSDVKELALDNRSGKLEALTDFEEFLSKINGLTSISDL 60
DB 1 MDMKRIHLELRNRTPAAVRELVLDNRCKNDGKIEGLTAEFVNLEFLSINVGLISVNL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSLTIPLKQLENKSLDL 116
DB 61 PKLPKLKLELSNRFGLDMLAEKLPNLTHLNLGNKIKDLSLTIPLKLEKLSL 120
QY 117 FNCVNTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDEEGEH-EEY 175
DB 121 FNCVNTNLNDYRESVFKLLPQLTYLDGYDREDQEAQPSDAE--VDSVDEEEDGEDEE 178
QY 176 DEDAQVVEDEGEDEE---EEGEEDVSG-----GDEDEEGYNDGEVDGEDEEELG 225
DB 179 DED-----DEGEDEEEDDEDDVDEGDEDDDEVESEEEFGLDEDEDEDEE-E 232
QY 226 EEERGGKRR 234
DB 233 EGGKGGKRR 241

RESULT 10

US-09-262-610-3
; Sequence 3, Application US/09262610
; Patent No. 6428949
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/262,610
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/766,738
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0177 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1498225
; US-09-262-610-3

Query Match
Best Local Similarity 56.2%; Score 683.5; DB 4; Length 251;
Matches 152; Conservative 34; Mismatches 40; Indels 23; Gaps 8;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGNECKLEALTDEFELEFLSKINGLTSIDL 60
DB 1 MDKRIHLELRNRTPAAVRELVDNCKSNCKIGLTAEFVNLFLSLINVLISVNL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLHSGNKIKDLSITIEPLKQLENKSLDL 116
DB 61 PKLPKLLKLELSENRIFFGLDMLAEKLPNLTNLHNSGNKLDISTIEPLKLECLKSLDL 120
QY 117 FNCVNTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSIEDHVEGLDDEEGEHEEY 175
DB 121 FNCVNTNLNDYRESVFKLLPQLTYLDGYDREDQAPDSAE--VDGVDEEEDDEGEDEE 178
QY 176 DEDAQVVEDEGEDEE---EEGEEDYSG-----GDEDEEGYNDGEVDGEDEEELG 225
DB 179 DED-----DEGDEEEDDEDEDEDEVDGEDEEVEEVEEFGLEDDEDEDEE-E 232
QY 226 EEERGQKRK 234
DB 233 EGGKGEKRK 241

RESULT 11
US-08-766-738-1
; Sequence 1, Application US/08766738
; Patent No. 5916749
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,738
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0177 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;

; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: 1813361
; US-08-766-738-1

Query Match
Best Local Similarity 55.6%; Score 676.5; DB 2; Length 251;
Matches 151; Conservative 34; Mismatches 41; Indels 23; Gaps 8;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGNECKLEALTDEFELEFLSKINGLTSIDL 60
DB 1 MDKRIHLELRNRTPAAVRELVDNCKSNCKIGLTAEFVNLFLSLINVLISVNL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLHSGNKIKDLSITIEPLKQLENKSLDL 116
DB 61 PKLPKLLKLELSENRIFFGLDMLAEKLPNLTNLHNSGNKLDISTIEPLKLECLKSLDL 120
QY 117 FNCVNTNLNDYGENVFKLLQLTYLDSCYWDHKEAPYSIEDHVEGLDDEEGEHEEY 175
DB 121 FNCVNTNLNDYRESVFKLLPQLTYLDGYDREDQAPDSAE--VDGVDEEEDDEGEDEE 178
QY 176 DEDAQVVEDEGEDEE---EEGEEDYSG-----GDEDEEGYNDGEVDGEDEEELG 225
DB 179 DED-----DEGDEEEDDEDEDEDEVDGEDEEVEEVEEFGLEDDEDEDEE-E 232
QY 226 EEERGQKRK 234
DB 233 EGGKGEKRK 241

RESULT 12
US-09-262-610-1
; Sequence 1, Application US/09262610
; Patent No. 6428949
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/262,610
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/766,738
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0177 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;

```


IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 1813361
US-09-262-610-1

Query Match 55.6%; Score 676.5; DB 4; Length 251;
Best Local Similarity 60.6%; Pred. No. 4.1e-57;
Matches 151; Conservative 34; Mismatches 41; Indels 23; Gaps 8;
QY 1 MEMGRIHSELRNAPSVDVKELALDNRSGNKGKLEALTDFEELEFYSKINGGLTSIDL 60
DQ 1 MDKRRRIHLELRNTPAARVRLVLDNCKSNCKIEGLTAETVNLNGLISVSNL 60
QY 61 PKL-KLRLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKQLENKSLDL 116
DQ 61 PKLKLKLELSENRIFGLDMLAEKLPNLTHLNLGNKLDLSTIEPLKLEKLSLDL 120
QY 117 FNCVTNLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDIEDHVEGLD-DEEGEHEPEEY 175
DQ 121 FNCVTNLNDYRESVFKLLPQLTYLDGYDRDQAPSDAE--VDGVDXXEEDGEGED 178
QY 176 DEDAQVDEGEDEE---EEGEEEDVSG-----GDEDEEGYNDGVDGEDEEELG 225
DQ 179 DED-----DEGEDEEEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEE-E 232
QY 226 EEEGQKRR 234
DQ 233 EGGGKRR 241

RESULT 13
US-08-466-603-5
; Sequence 5, Application US/08466603
; Patent No. 5726018
; GENERAL INFORMATION:
; APPLICANT: Pasternack, Gary R.
; APPLICANT: Kujajda, Francis P.
; TITLE OF INVENTION: No. 5726018el Mammalian Protein Associated With
; TITLE OF INVENTION: Uncontrolled Cell Division
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,603
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,503
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske Esq., Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 1107.47218
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9153
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-466-603-5

Query Match 53.3%; Score 648.5; DB 1; Length 182;
Best Local Similarity 77.5%; Pred. No. 1.2e-54;
Matches 134; Conservative 11; Mismatches 23; Indels 5; Gaps 2;
QY 65 LRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKQLENKSLDLFNC 121
DQ 1 VKKLESENKRISGDLEVLAEKCPNLKHLNLSGNKIKDLSTIEPLKLEKLSLDL 60
QY 122 TNLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDIEDHVEGLDDEEGEHEPEEYDQ 181
DQ 61 TNLNAYRENVFKLLPQWYLDGYDRNKEAPSDVGEYVE--DDDEDEDEEYDEY 118
QY 182 VEDEGEDEEEDVSGDEDEDEEGYNDGVDGEDEEELGEEERQKRR 234
DQ 119 VEDEGEDEEEDVSGDEDEDEEGYNDGVDGEDEEELGEEERQKRR 171

RESULT 14
US-08-314-503A-5
; Sequence 5, Application US/08314503A
; Patent No. 5734022
; GENERAL INFORMATION:
; APPLICANT: Pasternack, Gary R.
; APPLICANT: Kujajda, Francis P.
; TITLE OF INVENTION: No. 5734022el Mammalian Protein Associated With
; TITLE OF INVENTION: Uncontrolled Cell Division
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/314,503A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske Esq., Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 1107.47218
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9153
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-314-503A-5

Query Match 53.3%; Score 648.5; DB 1; Length 182;
Best Local Similarity 77.5%; Pred. No. 1.2e-54;
Matches 134; Conservative 11; Mismatches 23; Indels 5; Gaps 2;
QY 65 LRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKQLENKSLDLFNC 121
DQ 1 VKKLESENKRISGDLEVLAEKCPNLKHLNLSGNKIKDLSTIEPLKLEKLSLDL 60
QY 122 TNLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDIEDHVEGLDDEEGEHEPEEYDQ 181
DQ 61 TNLNAYRENVFKLLPQWYLDGYDRNKEAPSDVGEYVE--DDDEDEDEEYDEY 118
QY 182 VEDEGEDEEEDVSGDEDEDEEGYNDGVDGEDEEELGEEERQKRR 234
DQ 119 VEDEGEDEEEDVSGDEDEDEEGYNDGVDGEDEEELGEEERQKRR 171

Db 119 VEDEEEVEEERGEDVSGEEDEDEEGYNDGEVDDEDEEAGEEGSQRRK 171

RESULT 15
US-08-468-066-5
; Sequence 5, Application US/08468066
; Patent No. 5756676
; GENERAL INFORMATION:
; APPLICANT: Pasternack, Gary R.
; APPLICANT: KubaJda, Francis P.
; TITLE OF INVENTION: No. 5756676el Mammalian Protein Associated With
; TITLE OF INVENTION: Uncontrolled Cell Division
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,066
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,503
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske Esq., Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 1107.47218
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9153
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-066-5

Query Match 53.3%; Score 648.5; DB 1; Length 182;
Best Local Similarity 77.5%; Pred. No. 1.2e-54;
Matches 134; Conservative 11; Mismatches 23; Indels 5; Gaps 2;

QY 65 LRKLEL--RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPIKQLENLKSLLDFNCEV 121
Db 1 VRKLELSERISGDLEVLAEKCPNLKHLNLSGNKIKDLSTIEPIKLENLKSLLDFNCEV 60

QY 122 TNLNDYGENVFILLQLTYLDSCYWDHKEAPYSDIEDHVEGLDDEEGEHEEYEDDAQV 181
Db 61 TNLNAYRENVKLLPQVMYLDGYDRDNKEAPDSVEGYVE--DDDEDEDEEYDEYQAL 118

QY 182 VEDEGEVEEERGEDVSGDEEDEEGYNDGEVDGEDEELGEERGQRRK 234
Db 119 VEDEEEVEEERGEDVSGEEDEDEEGYNDGEVDDEDEEAGEEGSQRRK 171

Search completed: July 7, 2003, 14:59:31
Job time : 36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:56:03 ; Search time 30 Seconds
(without alignments)
896.945 Million cell updates/sec

Title: US-09-591-500A-4

Perfect score: 1216

Sequence: 1 MEMGRIHSELNRAPSDVK.....VDGEDEBELGEEGRGQKR 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pdb.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pdb.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pdb.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pdb.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pdb.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pdb.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pdb.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pdb.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pdb.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pdb.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pdb.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	82.3	249	9	US-10-213-700-4
2	683.5	56.2	251	9	US-10-213-700-3
3	676.5	55.6	251	9	US-10-213-700-1
4	175	14.4	76	9	US-10-101-487-36
5	173	14.2	180	9	US-10-101-487-116
6	173	14.2	197	9	US-10-101-487-51
7	171	14.1	197	9	US-10-101-487-114
8	170	14.0	200	9	US-10-101-487-53
9	169	13.9	179	9	US-10-101-487-107
10	169	13.9	181	9	US-10-101-487-45
11	166	13.7	176	9	US-10-101-487-70
12	166	13.7	177	9	US-10-101-487-48
13	166	13.7	177	9	US-10-101-487-115
14	166	13.7	179	9	US-10-101-487-46
15	166	13.7	186	9	US-10-101-487-44
16	166	13.7	187	9	US-10-101-487-50
17	166	13.7	191	9	US-10-101-487-81
18	166	13.7	198	9	US-10-101-487-42
19	165	13.6	174	9	US-10-101-487-72

165 13.6 175 9 US-10-101-487-57
20 165 13.6 176 9 US-10-101-487-56
21 165 13.6 240 9 US-10-101-487-75
22 165 13.6 350 9 US-10-101-487-58
23 160 13.2 387 9 US-09-893-519A-42
24 156 12.8 376 9 US-10-156-761-9889
25 150.5 12.4 740 9 US-10-128-714-3032
26 150.5 12.4 740 9 US-10-128-714-3032
27 147.5 12.1 119 10 US-09-864-761-44441
28 147.5 12.1 114 9 US-09-864-761-36168
29 146.5 12.0 440 9 US-10-025-380-1059
30 146.5 12.0 440 10 US-09-922-217-1059
31 146.5 12.0 440 10 US-09-833-263-1059
32 144.5 11.9 272 10 US-09-922-261-186
33 142.5 11.7 62 10 US-09-864-761-35828
34 141.5 11.6 89 9 US-10-101-487-43
35 141.5 11.6 89 10 US-09-864-761-44120
36 140.5 11.6 714 10 US-09-978-242-3
37 140.5 11.2 231 10 US-09-922-261-194
38 136.5 11.2 232 10 US-09-922-261-192
39 136.5 11.2 238 10 US-09-922-261-190
40 136.5 11.2 257 10 US-09-922-261-188
41 135.5 11.1 743 9 US-10-087-464-53
42 133 10.9 534 9 US-09-893-519A-37
43 133 10.9 1504 9 US-09-932-145-7
44 131.5 10.8 376 9 US-10-156-761-9888
45 131.5 10.8

ALIGNMENTS

RESULT 1

US-10-213-700-4

; Sequence 4, Application US/10213700

; Publication No. US20030022332A1

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/213,700

; FILING DATE: 06-Aug-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/766,738

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0177 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 249 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

Sequence 57, Appl
Sequence 56, Appl
Sequence 75, Appl
Sequence 58, Appl
Sequence 42, Appl
Sequence 9889, Ap
Sequence 3032, Ap
Sequence 8032, Ap
Sequence 4444, A
Sequence 36168, A
Sequence 1059, Ap
Sequence 1059, Ap
Sequence 1059, Ap
Sequence 186, App
Sequence 35828, A
Sequence 43, Appl
Sequence 4120, A
Sequence 3, Appli
Sequence 194, App
Sequence 192, App
Sequence 190, App
Sequence 188, App
Sequence 53, Appl
Sequence 37, Appl
Sequence 7, Appli
Sequence 9888, Ap

IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 403007
SEQUENCE DESCRIPTION: SEQ ID NO: 4;
US-10-213-700-4

Query Match 82.3%; Score 1001; DB 9; Length 249;
Best Local Similarity 86.1%; Pred. No. 2.4e-65;
Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;

QY 1 MEMGRIHLELRNAPSVDKVELALDNRSGKLEALTDREFEFLSKINGLTSIDL 60
DB 1 MEMGRIHLELRNAPSVDKVELALDNRSGKLEALTDREFEFLSKINGLTSIDL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLSCGNKIKDLSTIEPLKLENLKSIDL 116
DB 61 PKLKLKLELSENRFGLDMLAEKLPNLTLYLSCGNKIKDLSTIEPLKLECLKSLDL 120
QY 117 FNCVNTLNNDYGENVFKLLQLTYLDSCYWDHKEAPYSIEDHVEGLDDEGEHEEYD 176
DB 121 FNCVNTLNNDYGENVFKLLQLTYLDGYDRDKEAPDSAEYVGLDDEDEEYD 180
QY 177 EDAQVVEDEGEDEEEDVSGGDEDEEGYNDGEVDGDEDEELGEEERGGK 234
DB 181 EDAQVVEDEGEDEEEDVSGGDEDEEGYNDGEVDGDEDEELGEEERGGK 238

RESULT 2
US-10-213-700-3
; Sequence 3, Application US/10213700
; Publication No. US20030022332A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/213,700
; FILING DATE: 06-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,738
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0177 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1498225
; SEQUENCE DESCRIPTION: SEQ ID NO: 3;

US-10-213-700-3
Query Match 56.2%; Score 683.5; DB 9; Length 251;
Best Local Similarity 61.0%; Pred. No. 2.4e-42;
Matches 152; Conservative 34; Mismatches 40; Indels 23; Gaps 8;

QY 1 MEMGRIHLELRNAPSVDKVELALDNRSGKLEALTDREFEFLSKINGLTSIDL 60
DB 1 MDMMRRHLELRNTPAARVRELYLDNCKSDNGKTEGLTAEFVNLEFLSLINVLISVNL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLSCGNKIKDLSTIEPLKLENLKSIDL 116
DB 61 PKLKLKLELSENRFGLDMLAEKLPNLTLYLSCGNKIKDLSTIEPLKLECLKSLDL 120
QY 117 FNCVNTLNNDYGENVFKLLQLTYLDSCYWDHKEAPYSIEDHVEGLDDEGEHEEY 175
DB 121 FNCVNTLNNDYGENVFKLLQLTYLDGYDRDKEAPDSAE--VDGVDEEEDGEEDEE 178
QY 176 DEDAQVVEDEGEDEE---EEGEEDVSG-----GDEDEEGYNDGEVDGDEDEELG 225
DB 179 DED-----DEGEDEEEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEE-E 232
QY 226 EEEERGQKRK 234
DB 233 EGGGKRRK 241

RESULT 3
US-10-213-700-1
; Sequence 1, Application US/10213700
; Publication No. US20030022332A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/213,700
; FILING DATE: 06-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,738
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0177 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: 1813361
; SEQUENCE DESCRIPTION: SEQ ID NO: 1;

RESULT 5
US-10-101-487-116
Sequence 116, Application US/10101487
; Patent No. US20020169125A1

	Query Match	14.28;	Score 173;	DB 9;	Length 197;
	Best Local Similarity	42.08;	Pred. No. 1.7e-05;		
	Matches	37; Conservative	18; Mismatches 32;	Indels	0; Gaps 0;
QY	147 DKKAPSYDIEDHVGLDDEEGEEHESEYDEDQAQVDEDEGEFEEDVEESGGDEED	206			
Dd	70 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE	129			

OTHER INFORMATION: protein	
US-10-101-487-53	
Query Match	14.0%; Score 170; DB 9; Length 200;
Best Local Similarity	44.2%; Pred. No. 2.9e-05;
Matches	38; Conservative 16; Mismatches 32; Indels 0; Gaps 0;
RESULT 7	
US-10-101-487-114	
; Sequence 114, Application US/10101487	
; Patent No. US20020169125A1	
; GENERAL INFORMATION:	
; APPLICANT: LEUNG, DAVID W.	
; APPLICANT: BERGMAN, PHILIP A.	
; APPLICANT: LOFOQUIST, ALAN	
; APPLICANT: PIETZ, GREGORY E.	
; APPLICANT: TOMPKINS, CHRISTOPHER K.	
; APPLICANT: WAGGONER JR., DAVID W.	
; TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES	
; TITLE OF INVENTION: THEREOF	
; FILE REFERENCE: 077319/0329	
; CURRENT APPLICATION NUMBER: US/10/101,487	
; CURRENT FILING DATE: 2002-03-20	
; PRIOR APPLICATION NUMBER: 60/277,705	
; PRIOR FILING DATE: 2001-03-21	
; NUMBER OF SEQ ID NOS: 116	
; SOFTWARE: PatentIn Ver. 2.1	
; SEQ ID NO 114	
; TYPE: PRT	
; LENGTH: 197	
; ORGANISM: Artificial Sequence	
; FEATURE:	
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide	
US-10-101-487-114	
Query Match	14.1%; Score 171; DB 9; Length 197;
Best Local Similarity	39.8%; Pred. No. 2.4e-05;
Matches	37; Conservative 21; Mismatches 35; Indels 0; Gaps 0;
RESULT 8	
US-10-101-487-53	
; Sequence 53, Application US/10101487	
; Patent No. US20020169125A1	
; GENERAL INFORMATION:	
; APPLICANT: LEUNG, DAVID W.	
; APPLICANT: BERGMAN, PHILIP A.	
; APPLICANT: LOFOQUIST, ALAN	
; APPLICANT: PIETZ, GREGORY E.	
; APPLICANT: TOMPKINS, CHRISTOPHER K.	
; APPLICANT: WAGGONER JR., DAVID W.	
; TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES	
; TITLE OF INVENTION: THEREOF	
; FILE REFERENCE: 077319/0329	
; CURRENT APPLICATION NUMBER: US/10/101,487	
; CURRENT FILING DATE: 2002-03-20	
; PRIOR APPLICATION NUMBER: 60/277,705	
; PRIOR FILING DATE: 2001-03-21	
; NUMBER OF SEQ ID NOS: 116	
; SOFTWARE: PatentIn Ver. 2.1	
; SEQ ID NO 53	
; LENGTH: 200	
; TYPE: PRT	
; ORGANISM: Artificial Sequence	
; FEATURE:	
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic fusion	
OTHER INFORMATION: protein	
US-10-101-487-53	
Query Match	14.0%; Score 170; DB 9; Length 200;
Best Local Similarity	44.2%; Pred. No. 2.9e-05;
Matches	38; Conservative 16; Mismatches 32; Indels 0; Gaps 0;
RESULT 9	
US-10-101-487-107	
; Sequence 107, Application US/10101487	
; Patent No. US20020169125A1	
; GENERAL INFORMATION:	
; APPLICANT: LEUNG, DAVID W.	
; APPLICANT: BERGMAN, PHILIP A.	
; APPLICANT: LOFOQUIST, ALAN	
; APPLICANT: PIETZ, GREGORY E.	
; APPLICANT: TOMPKINS, CHRISTOPHER K.	
; APPLICANT: WAGGONER JR., DAVID W.	
; TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES	
; TITLE OF INVENTION: THEREOF	
; FILE REFERENCE: 077319/0329	
; CURRENT APPLICATION NUMBER: US/10/101,487	
; CURRENT FILING DATE: 2002-03-20	
; PRIOR APPLICATION NUMBER: 60/277,705	
; PRIOR FILING DATE: 2001-03-21	
; NUMBER OF SEQ ID NOS: 116	
; SOFTWARE: PatentIn Ver. 2.1	
; SEQ ID NO 107	
; LENGTH: 179	
; TYPE: PRT	
; ORGANISM: Artificial Sequence	
; FEATURE:	
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide	
US-10-101-487-107	
Query Match	13.9%; Score 169; DB 9; Length 179;
Best Local Similarity	41.1%; Pred. No. 3e-05;
Matches	37; Conservative 19; Mismatches 34; Indels 0; Gaps 0;
RESULT 10	
US-10-101-487-45	
; Sequence 45, Application US/10101487	
; Patent No. US20020169125A1	
; GENERAL INFORMATION:	
; APPLICANT: LEUNG, DAVID W.	
; APPLICANT: BERGMAN, PHILIP A.	
; APPLICANT: LOFOQUIST, ALAN	
; APPLICANT: PIETZ, GREGORY E.	
; APPLICANT: TOMPKINS, CHRISTOPHER K.	
; APPLICANT: WAGGONER JR., DAVID W.	
; TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES	
; TITLE OF INVENTION: THEREOF	
; FILE REFERENCE: 077319/0329	
; CURRENT APPLICATION NUMBER: US/10/101,487	
; CURRENT FILING DATE: 2002-03-20	
OTHER INFORMATION: protein	
US-10-101-487-53	
Query Match	14.0%; Score 170; DB 9; Length 200;
Best Local Similarity	44.2%; Pred. No. 2.9e-05;
Matches	38; Conservative 16; Mismatches 32; Indels 0; Gaps 0;
RESULT 11	
US-10-101-487-114	
; Sequence 114, Application US/10101487	
; Patent No. US20020169125A1	
; GENERAL INFORMATION:	
; APPLICANT: LEUNG, DAVID W.	
; APPLICANT: BERGMAN, PHILIP A.	
; APPLICANT: LOFOQUIST, ALAN	
; APPLICANT: PIETZ, GREGORY E.	
; APPLICANT: TOMPKINS, CHRISTOPHER K.	
; APPLICANT: WAGGONER JR., DAVID W.	
; TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES	
; TITLE OF INVENTION: THEREOF	
; FILE REFERENCE: 077319/0329	
; CURRENT APPLICATION NUMBER: US/10/101,487	
; CURRENT FILING DATE: 2002-03-20	
; PRIOR APPLICATION NUMBER: 60/277,705	
; PRIOR FILING DATE: 2001-03-21	
; NUMBER OF SEQ ID NOS: 116	
; SOFTWARE: PatentIn Ver. 2.1	
; SEQ ID NO 114	
; LENGTH: 197	
; TYPE: PRT	
; ORGANISM: Artificial Sequence	
; FEATURE:	
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide	
US-10-101-487-114	
Query Match	14.1%; Score 171; DB 9; Length 197;
Best Local Similarity	39.8%; Pred. No. 2.4e-05;
Matches	37; Conservative 21; Mismatches 35; Indels 0; Gaps 0;
RESULT 12	
US-10-101-487-53	
; Sequence 53, Application US/10101487	
; Patent No. US20020169125A1	
; GENERAL INFORMATION:	
; APPLICANT: LEUNG, DAVID W.	
; APPLICANT: BERGMAN, PHILIP A.	
; APPLICANT: LOFOQUIST, ALAN	
; APPLICANT: PIETZ, GREGORY E.	
; APPLICANT: TOMPKINS, CHRISTOPHER K.	
; APPLICANT: WAGGONER JR., DAVID W.	
; TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES	
; TITLE OF INVENTION: THEREOF	
; FILE REFERENCE: 077319/0329	
; CURRENT APPLICATION NUMBER: US/10/101,487	
; CURRENT FILING DATE: 2002-03-20	
; PRIOR APPLICATION NUMBER: 60/277,705	
; PRIOR FILING DATE: 2001-03-21	
; NUMBER OF SEQ ID NOS: 116	
; SOFTWARE: PatentIn Ver. 2.1	
; SEQ ID NO 53	
; LENGTH: 200	
; TYPE: PRT	
; ORGANISM: Artificial Sequence	
; FEATURE:	
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic fusion	

US-10-101-487-115

Query Watch 13.7%; Score 166; DB 9; Length 177;
Best Local Similarity 43.4%; Pred. No. 5e-05;
Matches 36; Conservative 17; Mismatches 30; Indels 0; Gaps 0;

QY 152 PYSDIEDHVEGLDDEEGEEHEEYDIDAQVVEDEGEFEFEFEEDVSGGDEDEGVN 211

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:54:17 ; Search time 40 seconds
(without alignments)
562.386 Million cell updates/sec

Title: US-09-591-500A-4
Perfect score: 1216
Sequence: 1 MEMGRRHSELNRAPSDVK.....VDGEDEELGEERQKRRK 234
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1001	82.3	249	2 S43309	probable HLA class
2	883	72.6	247	2 I59334	leucine-rich acid
3	667	54.9	272	2 JC7357	31K proliferation
4	526.5	43.3	134	2 S36375	PHAP1 protein - hu
5	294.5	24.2	447	2 T46146	hypothetical prote
6	243	20.0	229	2 T21714	hypothetical prote
7	230	18.9	225	2 T34456	hypothetical prote
8	177.5	14.6	556	2 B46024	neurofilament-L su
9	176	14.5	409	2 E86336	hypothetical prote
10	174.5	14.4	544	2 B44841	low molecularweig
11	170.5	14.0	792	2 T42963	hypothetical prote
12	167.5	13.8	590	2 A04037	glutamic acid-rich
13	166.5	13.7	543	1 QFMSL	hypothetical prote
14	165	13.6	721	2 S29795	hypothetical prote
15	165	13.6	797	2 A36811	probable cell divi
16	165	13.6	896	2 T24169	neurofilament tri
17	165	13.6	906	2 T24166	glutamic acid-rich
18	164	13.5	1019	2 T40813	related to actin-i
19	163.5	13.4	544	2 S07144	Ran GTPase activat
20	163	13.4	678	2 A34514	neurofilament tri
21	162.5	13.4	364	1 EDBESP	glutamic acid-rich
22	161	13.2	972	2 T49773	immediate-early pr
23	160	13.2	587	2 JC5300	related to actin-i
24	160	13.2	798	2 I50479	Ran GTPase activat
25	159.5	13.1	1300	2 T03166	neurofilament medi
26	158.5	13.0	636	2 T51893	probable immediate
27	158	13.0	279	2 S30766	related to Che-1 p
28	157	12.9	630	2 S29796	ASF1 protein - yea
29	156	12.8	81	2 T48398	hypothetical prote

30	156	12.8	1262	2 T33074	hypothetical prote
31	155	12.7	599	2 S18735	centromere protein
32	155	12.7	858	2 S15762	neurofilament trip
33	155	12.7	913	2 T52485	neurofilament prot
34	154	12.7	589	2 A36983	RNA1 homolog fagl
35	154	12.7	589	2 T52070	RNA1 protein homol
36	153	12.6	548	1 OFFGL	neurofilament trip
37	151	12.4	916	2 A27864	neurofilament trip
38	150	12.3	606	2 S70358	centromere protein
39	150	12.3	670	2 T28391	ORF MSV230 hypothe
40	149.5	12.3	407	1 EDBEQ3	immediate-early pr
41	148.5	12.2	262	2 T41536	yeast anti-silenci
42	148.5	12.2	1132	2 T43483	translation initia
43	147.5	12.1	168	2 F84499	hypothetical prote
44	147.5	12.1	845	2 A45669	neurofilament trip
45	147	12.1	325	2 T18283	hypothetical prote

ALIGNMENTS

RESULT 1
S43309
probable HLA class II-associated protein PHAP1 - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Nov-2000
C:Accession: S43309; S37222
R:Vaesen, M.; Barnikol-Watanabe, S.; Goetz, H.; Awml, L.A.; Cole, T.; Zimmermann, B.,
Biol. Chem. Hoppe-Seyler 375, 113-126, 1994
A:Title: Purification and characterization of two putative HLA class II associated pr
A:Reference number: S43309; MUID:94250340; PMID:8192856
A:Accession: S43309
A:Molecule type: mRNA
A:Residues: 1-249 <VAE>
A:Cross-references: EMBL:X75090; NID:q403006; PIDN:CAA52981.1; PID:q403007
C:Superfamily: Caenorhabditis elegans hypothetical protein T19H12.2
C:Keywords: signal transduction

Query Match 82.3%; Score 1001; DB 2; Length 249;
Best Local Similarity 86.1%; Pred. No. 2.7e-54;
Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;
QY 1 MEMGRRHSELNRAPSDVKELALDNSRSGKLEALTDDEEELEFLSKINGLTSDLS 60
Db 1 MEMGRRHLELRNTPSDVKELVDNSRSGKLEGLTDEFELEFLSTINVLGTSTANL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLSGNKIKDLSTIEPLKLENKLSLDL 116
Db 61 PKLNKAKLELSDNRVSGGLEVLAEKCPNLTLYLSGNKIKDLSTIEPLKLENKLSLDL 120
QY 117 FNCVETNLDYGENVFKLLQLTYLDSYWDHKEAPYSDIEDHVEGLDDEEGEHEBEYD 176
Db 121 FNCVETNLDYRENVFKLLPOLTYLDCYDRDKEAPSDAEGYVEGLDDEEDEDEYD 180
QY 177 EDAQVVEDEEEDDEEEDVSGGDEEBEGVNDGEVDEGEELGEERQKRRK 234
Db 181 EDAQVVEDEEEDDEEEDVSGGDEEBEGVNDGEVDEGEELGEERQKRRK 238

RESULT 2
I59334
leucine-rich acidic nuclear protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 17-Nov-2000
C:Accession: I59334
R:Matsuoka, K.; Taoka, M.; Satozawa, N.; Nakayama, H.; Ichimura, T.; Takahashi, N.; Y
Proc. Natl. Acad. Sci. U.S.A. 91, 9670-9674, 1994
A:Title: A nuclear factor containing the leucine-rich repeats expressed in murine cer
A:Reference number: I59334; MUID:95024022; PMID:7937870
A:Accession: I59334
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-247 <RES>

A:Cross-references: GB:D32209; NID:g511663; PIDN:BAA0908.1; PID:g511664
C:Superfamily: Caenorhabditis elegans hypothetical protein T19H12.2

Query Match 72.6%; Score 883; DB 2; Length 247;
Best Local Similarity 77.3%; Pred. No. 4.1e-47;
Matches 184; Conservative 16; Mismatches 32; Indels 6; Gaps 3;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGKLEALTDEFEELFLSKINGLTSIDL 60
Db 1 MEMDKRIYELNRTPSDVKELVLDNCRSIEGKIEGLTDEFEELFLSTINVLGTSISNL 60

QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKOLENLSIDL 116
Db 61 PLNKLUKLELSENRIKGLVLAELKCPNLKHLNLSGNKIKDLSTIEPLKOLENLSIDL 120

QY 117 FNCVNTLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDIEDHVEGLDDEEGHEEYD 176
Db 121 FNCVNTLNDYGENVFKLLPQVMYLDGYDRNKEAPSDVEGYVE--DDDEDEDEYD 178

QY 177 EDQVVEDEEGEEDEEEDVSGDEDEDEGYNDGEVDGEDEEELGEERQKRK 234
Db 179 EYALVDEDEEEDVEEEDVSGDEDEDEGYNDGEVDGEDEEEDAAEEEGSKRK 236

RESULT 3

JC7357
31K proliferation related acidic leucine-rich protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 17-Nov-2000
A:Accession: JC7357
R:Mutai, H.; Toyoshima, Y.; Sun, W.; Hattori, N.; Tanaka, S.; Shiota, K.
Biochem. Biophys. Res. Commun. 274, 427-433, 2000
A:Title: PAL31, a novel nuclear protein, expressed in the developing brain.
A:Reference number: JC7357
A:Accession: JC7357
A:Molecule type: mRNA
A:Residues: 1-272 <MUT>
A:Cross-references: DDBJ:AB025581
A:Gene: pal31
C:Superfamily: Caenorhabditis elegans hypothetical protein T19H12.2
C:Keywords: brain; nucleus; tandem repeat

Query Match 54.9%; Score 667; DB 2; Length 272;
Best Local Similarity 55.3%; Pred. No. 6.6e-34;
Matches 147; Conservative 39; Mismatches 45; Indels 34; Gaps 7;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGKLEALTDEFEELFLSKINGLTSIDL 60
Db 1 MDKRRTHLELRNRTPAAVQELVLDNCKANDGKIEGLTDFVNLFLSLINVLFSVSDL 60

QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKOLENLSIDL 116
Db 61 PKLKLKLELSENRIKGLVLAELKCPNLTHLYLSGNKIKDLSTIEPLKLRDCKLSIDL 120

QY 117 FNCVNTLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDIE-----DHVEGLDD 165
Db 121 FNCVNTLNDYGENVFKLLPQVMYLDGYDRNKEAPSDVEGYVE--DDDEDEDEYD 178

QY 166 EEEGE---HHEEYDEDAQVVEDEEGEEDEEEDVSGDEDEE-----BEGYNDGEV 215
Db 181 EEEGEDEDEDEEDG---EEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 237

QY 216 DGEDEEELGEE-----RGQKRK 234
Db 238 DEDEDEDEDEDEEESGAGK 262

RESULT 4

S36375
PHAPI protein - human (fragments)
C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Nov-2000
C:Accession: S36375
R:Vaesen, M.; Barnikol-Watanabe, S.; Goetz, H.; Kratzin, H.D.; Hilschmann, N.
submitted to the Protein Sequence Database, September 1993
A:Reference number: S36374
A:Accession: S36375
A:Molecule type: protein
A:Residues: 1-26;27-45;46-72;73-134 <VAE>
C:Superfamily: Caenorhabditis elegans hypothetical protein T19H12.2
C:Keywords: blocked amino end
F:/Modified site: blocked amino end (Met) #status experimental

Query Match 43.3%; Score 526.5; DB 2; Length 134;
Best Local Similarity 73.9%; Pred. No. 1.1e-25;
Matches 116; Conservative 1; Mismatches 17; Indels 23; Gaps 3;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGKLEALTDEFEELFLSKINGLTSIDL 60
Db 1 MEMGRIHSELNRAPSDVKELVLDN--SNEGKLEGLTDEFEELFLSDN-----49

QY 61 PKLKLKLELRYSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKOLENLSIDL 120
Db 50 -----RVSGGLEVLAEKCPNLTHLYLSG--IKDLSTIEPLKLENLSIDL 97

QY 121 VTNLNDYGENVFKLLQLTYLDSYWDHKEAPYSIDIE 157
Db 98 VTNLNDYGENVFKLLPQVMYLDGYDRDKEAPSDAE 134

RESULT 5

T46146
hypothetical protein T3A5.70 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46146
R:Blöcker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quétier, F.; S.
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23024
A:Accession: T46146
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <BLO>
A:Cross-references: EMBL:AL132979
A:Experimental source: cultivar Columbia; BAC clone T3A5
C:Genetics:
A:Map position: 3
A:Introns: 369/3
A:Note: T3A5.70

Query Match 24.2%; Score 294.5; DB 2; Length 447;
Best Local Similarity 34.1%; Pred. No. 5.6e-11;
Matches 86; Conservative 37; Mismatches 88; Indels 41; Gaps 8;

QY 20 KETALDNR--RSNEGKL--EALTDEFEELFLSKINGLTSIDLKPL-KLRKLEL---RV 72
Db 23 RTITLGVKCVGRLPFPVSVLEKFNQLHLSVANIGVSSLEQFPRLGNLKLSDNRI 82

QY 73 SGGLEVLAEK-CPNLTHLYLSGNKIKDLSTIEPLKOLENLSIDLFCNCEVTLNDYGENV 131
Db 83 TVGLEFLVLAELGDFCDLSDNRIQFVEDLAPLAEK-LVSLDLYECPVTRKDYRSRV 141

QY 132 FKLLQLTYLDSYWDHKEAPYSIDIEDHVEGLDDEEGE-----170
Db 142 FGLIKTLKYLKDTADAGNERPESDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 201

QY 171 -----HHEEYDEDAQVVEDEEGEEDEEEDVSGDEDEEGYNDGEV-----219
Db 202 RMSNGHSRVDGVVDVEDEESDAEDESEQATGVNGTYSRANGFRLEAVNGEVRDDG 261

QY 220 DEEELGEERQ 231
Db 262 DDESSEGEVGE 273

```
RESULT 6
T21714
hypothetical protein F33H2.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000
C:Accession: T21714
R:Cottage, A.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19463
A:Accession: T21714
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-229 <WIL>
A:Cross-references: EMBL:Z81526; PIDN:CA804265.1; GSPDB:GN00019; CESP:F33H2.3
A:Experimental source: clone F33H2
C:Genetics:
A:Gene: CESP:F33H2.3
A:Map position: 1
A:Introns: 199/3
C:Superfamily: Caenorhabditis elegans hypothetical protein T19H12.2

Query Match 20.0%; Score 243; DB 2; Length 229;
Best Local Similarity 31.9%; Pred. No. 3.7e-08;
Matches 75; Conservative 40; Mismatches 88; Indels 32; Gaps 9;

QY 10 ELRRAPSDYKELALDNRSECKLEALTDEFEFEFLSKINGLTSISDLPR-----L 63
DB 11 ELRRDPATVDTFLDN--AEDGQIGGLTDLINLEMLSMVKGLTTLAGFPTLPALTYL 68
QY 64 KRLRLRVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKLENKLSLDLF-NCEVT 122
DB 69 DISDNQLGDNASDFVLVKNAPDLKKITLANKL-SLDNLRCLKVLNPLFELDLSNPISG 127
QY 123 NLNDYGENVFKLLQLTYLDSYWDHKEAPYSIEDHV--EGLDDEEGSEHEE-----Y 175
DB 128 LLEDYRKFMFEMIPSLKILDCDVGDEE----VEEFAGEGGSESGDGEDGPGLSY 182
QY 176 DEDAQVDEDEGESEEEEDVSGGDEDEEGYNDGEVDGEE--DEEELGEE 228
DB 183 LEKSQSDDETDYAPE-----GGDAEPRTGKRGASDNGEEDPNKKAAGDDE 229

RESULT 7
T34456
hypothetical protein T19H12.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 08-Sep-2000
C:Accession: T34456
R:Davidson, S.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid T19H12.
A:Reference number: Z21528
A:Accession: T34456
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-225 <DAV>
A:Cross-references: EMBL:U97009; PIDN:AAC69027.1; GSPDB:GN00023; CESP:T19H12.2
A:Experimental source: strain Bristol N2; clone T19H12
C:Genetics:
A:Gene: CESP:T19H12.2
A:Map position: 5
A:Introns: 41/3; 199/3
C:Superfamily: Caenorhabditis elegans hypothetical protein T19H12.2

Query Match 18.9%; Score 230; DB 2; Length 225;
Best Local Similarity 30.0%; Pred. No. 2.3e-07;
Matches 71; Conservative 34; Mismatches 86; Indels 46; Gaps 7;

QY 9 SELRRAPSDYKELALDNRSECKLEALTDEFEFEFLSKINGLTSISDLP-----K 62
DB 7 SELRGREPVTDTFLDNTQG--GVIGGINEKLTKELLSMVKGLTTLKGMPLPALNY 64
```

```
QY 63 LXLRLKLRVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKLENKLSLDLF-NCEV 121
DB 65 LLDLSELGDDASDFVLKCAPEIKKILTSGNRL-TLDNVRTLKMLPNLMELDLNNSSL 123
QY 122 TNLNDYGENVFKLLQLTYLDSYWDHKEAPYSIEDHVGLDDDEEGEHEEYDEDAQV 181
DB 124 GLDDYRVKMFEMIPSLKILDCG-----DVGDEE 152
QY 182 VEDE----EGEEEGE--EDVSGGDEDEEGYNDGEVDGEDEEELGEEERGGKR 233
DB 153 VEEFAAGCAEDSDGSDDEGPGLSYLNKSQFSDDETDYVVPVPEAGDAETRGAKR 209

RESULT 8
B46024
neurofilament-L subunit - quail
C:Species: Coturnix coturnix (quail)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
C:Accession: B46024
R:Ohara, O.; Gahara, Y.; Miyake, T.; Teraoka, H.; Kitamura, T.
J. Cell Biol. 121, 387-395, 1993
A:title: Neurofilament deficiency in quail caused by nonsense mutation in neurofilament
A:Reference number: A46024; MUID:93224534; PMID:8468353
A:Accession: B46024
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-556 <OHA>
A:Experimental source: subsp. japonica, TKP
A>Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:129455, NCBIP:129482)
C:Superfamily: cytoskeletal keratin

Query Match 14.6%; Score 177.5; DB 2; Length 556;
Best Local Similarity 26.5%; Pred. No. 0.00096;
Matches 65; Conservative 38; Mismatches 91; Indels 51; Gaps 9;

QY 20 KELALDNRSECKLEALTDEFEFEFLSKINGLTSISDLPLKLRKLELRVSGGLEVL 79
DB 317 KYLEIATR--GMNEALEKQLELE--EKQSANISALQD----YINKLENELRTTKSM 367
QY 80 AEKCPNLTHLYLSGNKIKDLSTIEPLKLENKLSLDLFNCEVTNUN----- 125
DB 368 AR-----YL--KEYQLLNVMKALDIEIAAYRKLLEGEETRLSFTSVGVTSGYTQT 417
QY 126 --DYGENVFKLLQLTYL-----DSCYWDHKEAPYSIEDHVGLD-----DEEGE 170
DB 418 APTFGRSAYSGLQSTSYLMTTRSFYPSYSHVQEOEIEETIEAKAGEAKAAPAEGE 477
QY 171 HEYEYDEDAQVVEDEGESEEEEDVSGGDEDEEGYNDGEVDGEDE--BELGEE 227
DB 478 EEKEEGEAEAGGEAEAESEEGEAESEEGEAESEEGEAESEEGEAESEEGEAESE 537
QY 228 ERGQK 232
DB 538 EKEEK 542

RESULT 9
E86336
hypothetical protein F14O10.11 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: E86336
R:theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon,
Chen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
```

A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86336
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <STO>
A:Cross-references: GB:AE005172; NID:g9558597; PIDN:AAF88160.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 14.5%; Score 176; DB 2; Length 409;
Best Local Similarity 25.4%; Pred. No. 0.00085;
Matches 68; Conservative 41; Mismatches 91; Indels 68; Gaps 7;
QY 2 EMGRRTHSELNRNAPSVDKELALDNRSGKLEALTFEELEFLSKINGLTSIS--- 58
Db 129 EQARLIPSSFHQVDP-DMKLLAADKINR-----LTLISGPKETISVYT 172
QY 59 --DLPLKLRKLELRVSGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKQLENKSLDL 116
Db 173 LVGNPKWTFRK-----NENPKWALDKRIKNVSMGSMFLWNVYVDG 215
QY 117 FNCEVTNLNDYGE-----NVFKLLQLTYL---DSCVW----- 146
Db 216 KRLVREMKDILIEGVVHYDMEANSRWVLCSTKWNPDSDFYKTPSMVHVGEITVOA 275
QY 147 DHKEAPYSIEDHVEGLDDEEHEEVEDEDAQVVEDEEHEEVEEVEEVEEVEEVEE 206
Db 276 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE 335
QY 207 EEGYNDGEVDGEDEELGEEERGOKR 234
Db 336 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE 363

RESULT 10

B44841
Low molecular weight neurofilament protein XNF-L - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
R:Accession: B44841
R:Charnas, L.R.; Szaro, B.G.; Gainer, H.
J. Neurosci. 12, 3010-3024, 1992
A:Title: Identification and developmental expression of a novel low molecular weight neurofilament protein
A:Reference number: A44841; MUID:92356194; PMID:149494
A:Accession: B44841
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-544 <CHA>
A:Experimental source: brain
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBI:110225, NCBI:110226)
C:Superfamily: cytoskeletal keratin

Query Match 14.4%; Score 174.5; DB 2; Length 544;
Best Local Similarity 27.6%; Pred. No. 0.0014;
Matches 72; Conservative 34; Mismatches 96; Indels 59; Gaps 11;
QY 13 NRAPSDVKELALDNRSGKLEALTFEELEFLSKINGLTSISDLPLKLRKLELR 72
Db 302 SRRLMSAKGLETEACK--GVNEALQROIQELE--DKQSGETAGMD---AINKLEEL 352
QY 73 SGGLEVLAEKCPNLTHLYLSGNKIKDLSTIEPLKQLENKSLDLFNCVNTLN----- 125
Db 353 RNTKSEMAR-----YL--KEYQDLLNVKMLDIEAAYRKLLEGSETRLSFSGVGI 402
QY 126 -----DYGENFKLLQLTYLDS-----CYWDHKEAPYSIEDHVEGLDDEE--E 170
Db 403 TSGYTQSPVFGRSAYS--LQSSYMTSRAFTYTSVSHVQEQLDIETIESRAEAKAE 461
QY 171 HEEVEYDEDAQVVEDEEHEEVEE-----EEEDVSGGDEEDBEG-----YNDG 213
Db 462 APEEEEEEAAEEEGEGEABEGEGEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAE 521

QY 214 EVDGEDEEEELGEEERGOKR 234
Db 522 EEEGSGKDEAAEESEKKEK 542

RESULT 11

T42963
Hypothetical protein 48 - ateline herpesvirus 3 (strain 73)
C:Species: ateline herpesvirus 3
A:Variety: strain 73
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T42963
R:Albrecht, J.C.; Fleckenstein, B.
A:Description: Primary structure of the herpesvirus ateles genome.
A:Reference number: 222274
A:Accession: T42963
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-792 <ALB>
A:Cross-references: EMBL:AF083424; PIDN:AAC95573.1
A:Experimental source: strain 73

Query Match 14.0%; Score 170.5; DB 2; Length 792;
Best Local Similarity 28.7%; Pred. No. 0.0038;
Matches 60; Conservative 37; Mismatches 69; Indels 43; Gaps 7;
QY 26 NSRNECKLALTFEELEFLSKINGLTSISDLPLKLRKLELRVSGGLEVLAEKCPN 85
Db 378 SGRNNKYKGANDKXSID--KNESEGG--DHSEINREKNRK-----RKKPN 421
QY 86 LTHLYLSGNKI--KDLSTIEPLKQLENKSLDLFNCVNTLNNDYGENVFKLLQLTYLDS 143
Db 422 -----GFRVGDKVEGEEKSVKSGGKKS--EKDSEEAEDKDEEN----- 460
QY 144 CYWDHKEAPYSIEDHVEGLDDEEHEEVEDEDAQVVEDEEHEEVEEVEEVEEVEE 203
Db 461 ----KKKGDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 516
QY 204 EEDEGYNDGEVDGEDEELGEEERGOKR 232
Db 517 EEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 545

RESULT 12

A40437
glutamic acid-rich protein, retinal - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 05-Nov-1999
R:Accession: A40437
R:Sugimoto, Y.; Yatsunami, K.; Tsujimoto, M.; Khorana, H.G.; Ichikawa, A.
Proc. Natl. Acad. Sci. U.S.A. 88, 3116-3119, 1991
A:Title: The amino acid sequence of a glutamic acid-rich protein from bovine retina
A:Reference number: A40437; MUID:91195303; PMID:2014230
A:Accession: A40437
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-590 <SDG>
A:Cross-references: GB:M61185; NID:g163077; PIDN:AAA30536.1; PID:g163078

Query Match 13.8%; Score 167.5; DB 2; Length 590;
Best Local Similarity 43.5%; Pred. No. 0.0042;
Matches 40; Conservative 18; Mismatches 29; Indels 5; Gaps 3;
QY 147 DHKEAPYSIEDHVEGLDDEEHEEVEDEDAQVVEDEEHEEVEEVEEVEEVEE 205
Db 359 DEEEKEDGEEEGEKEKEE--EGEKEEEEGEKEKEEKEKEKEKEKEKEKEKEKE 417
QY 206 DEEGYNDGEVDGEDEELG---EEERGOKR 234
Db 418 DEEGKEKEEGEKGEEGEGEKEGCKE 449

[illegible]

A:Gene: 48

```
Query Match      13.6%; Score 165; DB 2; Length 797;
Best Local Similarity 28.8%; Pred. No. 0.0082;
Matches 72; Conservative 37; Mismatches 89; Indels 52; Gaps 13;

QY 22 LAL--DNSRSNGKLEALTDPEFEFLSKINGGLTSTSDL-----PKLKRKLE 69
   ||| :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 239 LALHPNDKASYSNLIKFLTSNSHREHVTQKVNRAFMQSSLYKIIDTEKNPSPKTKMLM 298
   :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
QY 70 LAVSG-----GLEVLAE---KCPNLTHLYLSGNKIKDLSTIEPL--KOLENLKSIDLPLNC 119
   :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 299 ISILSGRGIGMDLFCQSQVLRKAPLIDH-----KLSPVSEYEDFDEDELCISDDEVDS 352
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 120 EVTN---LNDYGENVFKLL-OLTYLDCYWDHKEAPYSDIEDHVEGLDDEEGEHE--- 172
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 353 EDGNLCVLDDSESVNSVALRQVLTVDK---QANEKEYKKIIDKSDDRDRDKDEYELEN 409
   ||| ||| :| | :| | :| | :| | :| | :| | :| | :| | :| |
QY 173 EY--DEDAQVVEDEGE-EEEEEGEEDVSGGDEEDEEGYNDG-----EVDGE 218
   ||| ||| :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 410 EYNRDEEDEGEDEGEDEKEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDE 469
   ||| ||| :| | :| | :| | :| | :| | :| | :| | :| | :| |
QY 219 EDEEELGEE 228
Db 470 DDEDEGEDE 479
```

Search completed: July 7, 2003, 14:58:59
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:50:37 ; Search time 22 Seconds
(without alignments)
441.157 Million cell updates/sec

Title: US-09-591-500A-4

Perfect score: 1216

Sequence: 1 MEMGRIHSELNRAPSDVK.....VDGEDEEELGEERQKPK 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1001	82.3	249	1 PHAL_HUMAN	P39687 homo sapien
2	883	72.6	247	1 LANP_RAT	P49911 rattus norv
3	624.5	51.4	173	1 LANP_BOVIN	P51122 bos taurus
4	177.5	14.6	555	1 NFL_COTUA	Q02916 coturnix co
5	174.5	14.4	544	1 NFL_XENLA	P35616 xenopus lae
6	174	14.3	541	1 NFL_RAT	P19527 rattus norv
7	167.5	13.8	1394	1 CNG4_BOVIN	Q28181 bos taurus
8	166.5	13.7	542	1 NFL_MOUSE	P08551 mus musculu
9	165	13.6	721	1 YCP2_OENPI	P31568 oenothera p
10	165	13.6	737	1 VG48_HSVSA	Q01033 herpesvirus
11	163.5	13.4	543	1 NFL_HUMAN	P07196 homo sapien
12	163	13.4	678	1 GARP_PLAUF	P13816 plasmodium
13	162.5	13.4	364	1 IE68_PRVKA	P24827 pseudorabie
14	160	13.2	587	1 RGPI_HUMAN	P46060 homo sapien
15	158	13.0	279	1 ASPI_YEAST	P32447 saccharomyc
16	157	12.9	630	1 YCF2_OENVI	P31569 oenothera v
17	156	12.8	599	1 CENB_MOUSE	P27790 mus musculu
18	155	12.7	599	1 CENB_HUMAN	P07199 homo sapien
19	155	12.7	857	1 NFL_CHICK	P16053 gallus gall
20	154.5	12.7	548	1 NFL_PTG	P02547 sus scrofa
21	154.5	12.7	580	1 NFL_XENLA	O13066 xenopus lae
22	154	12.7	589	1 RGPI_MOUSE	P46061 mus musculu
23	153	12.6	539	1 Z173_HUMAN	Q12899 homo sapien
24	151	12.4	915	1 NFM_HUMAN	P07197 homo sapien
25	150	12.3	606	1 CENB_CRIGR	P48988 cricetus
26	149.5	12.3	407	1 IE68_HSVSA	Q01042 herpesvirus
27	149	12.3	239	1 CENB_SHEEP	P49451 ovis aries
28	148.5	12.2	1220	1 IF2P_HUMAN	O60841 mus musculu
29	147	12.1	554	1 NFL_BOVIN	P02548 bos taurus
30	147	12.1	845	1 NFM_RAT	P12839 rattus norv
31	146.5	12.0	848	1 NFM_MOUSE	P08553 mus musculu
32	146	12.0	743	1 ABRA_PLAFC	P22620 plasmodium
33	145.5	12.0	810	1 NFM_BOVIN	O77788 bos taurus

RESULT 1

PHAL_HUMAN

ID	PHAL_HUMAN	STANDARD	PRT	249 AA
AC	P39687			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Potent heat-stable protein phosphatase 2A inhibitor I1PP2A (HLA-DR associated protein I) (PHAPI) (Acidic nuclear phosphoprotein PP32)			
DE	(Cerebellar leucine rich acidic nuclear protein).			
GN	PHAPI OR LANP			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI-TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE=94250340; PubMed=8192856;			
RA	Vaesens M., Barnikol-Watanabe S., Goetz H., Adil Avni L., Cole T., Zimmermann B., Kratzin H.D., Hilschmann N.;			
RT	"Purification and characterization of two putative HLA class II associated proteins: PHAPI and PHAPII."			
RL	Biol. Chem. Hoppe-Seyler 375:113-126(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96240314; PubMed=8679524;			
RA	Li M., Makkinje A., Damuni Z.;			
RT	"Molecular identification of I1PP2A, a novel potent heat-stable inhibitor protein of protein phosphatase 2A.";			
RL	Biochemistry 35:6998-7002(1996).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Chen T.H., Brody J.R., Romantsev F.E., Yu J.G., Kayler A.E., Voneil E., Kuna Jda F.P., Pasternack G.R.;			
RL	Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98013170; PubMed=9353121;			
RA	Matilla A., Koshy B.T., Cummings C.J., Isobe T., Orr H.T., Zoghbi H.Y.;			
RT	"The cerebellar leucine-rich acidic nuclear protein interacts with ataxin-1.";			
RL	Nature 389:974-978(1997).			
RN	[5]			
RP	FUNCTION: INHIBITOR PROTEIN OF PROTEIN PHOSPHATASE 2A.			
CC	-!- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC, DIFFUSELY DISTRIBUTED IN THE CYTOSOL.			
CC	-!- PTM: PHOSPHORYLATED ON SERINE RESIDUES.			
CC	-!- PTM: THE N-TERMINUS IS BLOCKED.			
CC	-!- SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).			
CC	-!- SIMILARITY: HOMOLOG OF RAT LANP.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way			

34	145.5	12.0	3135	1	S230_PLAFO	Q08372 plasmodium
35	144.5	11.9	1035	1	CC68_YEAST	P32558 saccharomyc
36	144	11.8	593	1	SANT_PLA7	Q03400 plasmodium
37	144	11.8	802	1	NAB3_YEAST	P38996 saccharomyc
38	143	11.8	706	1	NUCL_MOUSE	P09405 mus musculu
39	142.5	11.7	411	1	MP62_LYTPI	P47245 lytechinus
40	142.5	11.6	1161	1	NRDC_RAT	P08199 mesocricetu
41	140.5	11.6	713	1	NUCL_MESAU	P19338 homo sapien
42	140	11.5	706	1	NUCL_HUMAN	P53115 saccharomyc
43	140	11.5	1489	1	YGP0_YEAST	P33441 saccharomyc
44	139.5	11.5	392	1	MFT1_YEAST	P13383 rattus norv
45	138.5	11.4	712	1	NUCL_RAT	

ALIGNMENTS

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X75090; CAA52981.1; -;
DR EMBL; U60823; AAC50570.1; -;
DR EMBL; U73477; AAB39706.1; -;
DR EMBL; AF025684; AAB91548.1; -;
DR PIR; S43309; S43309.
DR PIR; S37222; S37222.
DR Genew; HGNC:13233; ANP32A.
DR MIM; 600832; -;
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR Pfam; PF00560; LRR; 2.
DR SMART; SM00446; LRRcap; 1.
KW Phosphorylation; Leucine-rich repeat; Repeat; Nuclear protein.
FT REPEAT 44 63 LRR 1.
FT REPEAT 64 86 LRR 2.
FT REPEAT 87 111 LRR 3.
FT REPEAT 115 138 LRR 4.
FT DOMAIN 168 249 ASP/GLU-RICH (HIGHLY ACIDIC).
SQ SEQUENCE 249 AA; 28585 MW; CA2D1A756FBA04 CRC64;

Query Match 82.3%; Score 1001; DB 1; Length 249;
Best Local Similarity 86.1%; Pred. No. 1.4e-54;
Matches 205; Conservative 7; Mismatches 22; Indels 4; Gaps 2;
QY 1 MEMGRIHSELNRAPSDVKELALNSRNECKLEALTDDEFEELFLSKINGLTSIDL 60
DB 1 MEMGRIHSELNRAPSDVKELALNSRNECKLEALTDDEFEELFLSKINGLTSIDL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLSGNKIKDLSTIEPLKLENKLSLD 116
DB 61 PKLKLKLELSDNRYSVGGLEVLAEKCPNLTLYLSGNKIKDLSTIEPLKLENKLSLD 120
QY 117 FNCVTLNDYGNVFKLLQLTYLDSCYWDHKEAPYSIEDHVGLEDDEEHEEYD 176
DB 121 FNCVTLNDYGNVFKLLPOLTYLDGYDRDKAPSDAEGYVGGLEDDEEHEEYD 180
QY 177 EDAQVDEDEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEYD 234
DB 181 EDAQVDEDEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEYD 238

RESULT 2
LAMP_RAT LAMP_BOVIN STANDARD; PRT; 247 AA.
AC P49911;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucine-rich acidic nuclear protein.
GN LAMP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Wistar; TISSUE=Brain;
RX MEDLINE=95024022; PubMed=7937870;
RA Matsuo K., Taoka M., Satozawa N., Nakayama H., Ichimura T.,
RT Takahashi N., Yamakuni T., Song S.-Y., Isobe T.;
RT "A nuclear factor containing the leucine-rich repeats expressed in
RT murine cerebellar neurons";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9670-9674 (1994).
CC -!- FUNCTION: MAY PLAY A ROLE IN A SIGNAL TRANSDUCTION PATHWAY THAT
CC DIRECTS DIFFERENTIATION OF CEREBELLAR NEURONS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: WIDELY DISTRIBUTED IN THE CENTRAL NERVOUS
CC SYSTEM, WITH AN ABUNDANT EXPRESSION IN THE CEREBELLUM.
CC -!- DEVELOPMENTAL STAGE: IT IS MODERATELY EXPRESSED IN THE CEREBELLUM

CC ON POSTNATAL DAY 7 IN THE EXTERNAL GRANULE AND PERKINJE CELLS,
CC INCREASES IN THE SECOND POSTNATAL WEEK AND THEREAFTER DECREASES
CC TO AN ADULT LEVEL.
CC -!- SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).
CC -!- SIMILARITY: HOMOLOG OF HUMAN PHAP-1.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; D32209; BAA06908.1; -;
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR Pfam; PF00560; LRR; 2.
DR SMART; SM00446; LRRcap; 1.
KW Phosphorylation; Leucine-rich repeat; Repeat; Nuclear protein.
FT REPEAT 44 63 LRR 1.
FT REPEAT 64 86 LRR 2.
FT REPEAT 87 111 LRR 3.
FT REPEAT 115 138 LRR 4.
FT DOMAIN 165 247 ASP/GLU-RICH (HIGHLY ACIDIC).
SQ SEQUENCE 247 AA; 28564 MW; 0919739645557462 CRC64;

Query Match 72.6%; Score 883; DB 1; Length 247;
Best Local Similarity 77.3%; Pred. No. 2e-47;
Matches 184; Conservative 16; Mismatches 32; Indels 6; Gaps 3;
QY 1 MEMGRIHSELNRAPSDVKELALNSRNECKLEALTDDEFEELFLSKINGLTSIDL 60
DB 1 MEMGRIHSELNRAPSDVKELALNSRNECKLEALTDDEFEELFLSKINGLTSIDL 60
QY 61 PKL-KLRKLEL---RVSGGLEVLAEKCPNLTLYLSGNKIKDLSTIEPLKLENKLSLD 116
DB 61 PKLKLKLELSDNRYSVGGLEVLAEKCPNLTLYLSGNKIKDLSTIEPLKLENKLSLD 120
QY 117 FNCVTLNDYGNVFKLLQLTYLDSCYWDHKEAPYSIEDHVGLEDDEEHEEYD 176
DB 121 FNCVTLNDYGNVFKLLPOLTYLDGYDRDKAPSDVGEYV---DDDEDEDEEYD 178
QY 177 EDAQVDEDEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEYD 234
DB 179 EDAQVDEDEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEHEEYD 236

RESULT 3
LAMP_BOVIN LAMP_BOVIN STANDARD; PRT; 173 AA.
AC P51122;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucine-rich acidic nuclear protein (Fragments).
GN LAMP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=95024022; PubMed=7937870;
RA Matsuo K., Taoka M., Satozawa N., Nakayama H., Ichimura T.,
RT Takahashi N., Yamakuni T., Song S.-Y., Isobe T.;
RT "A nuclear factor containing the leucine-rich repeats expressed in
RT murine cerebellar neurons";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9670-9674 (1994).
CC -!- FUNCTION: MAY PLAY A ROLE IN A SIGNAL TRANSDUCTION PATHWAY THAT
CC DIRECTS DIFFERENTIATION OF CEREBELLAR NEURONS.

RT	weight neuronal intermediate filament protein expressed in xenopus laevis."	
RL	J. Neurosci. 12:3010-3024(1992).	
CC	- - FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M, AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.	
CC	- - MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM FILAMENTS.	
CC	- - SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
DR	EMBL; M86654; AAA03018.1; -.	
DR	PIR; B44841; B44841.	
DR	InterPro; IPR001664; IF.	
DR	Pfam; PF00038; filament; 1.	
DR	PROSITE; PS00226; IF; 1.	
KW	Intermediate filament; Coiled coil; Neurone.	
FT	DOMAIN 1 87	
FT	DOMAIN 88 390	
FT	DOMAIN 391 544	
FT	DOMAIN 545 720	
FT	DOMAIN 721 877	
FT	DOMAIN 878 1119	
FT	DOMAIN 1120 132	
FT	DOMAIN 133 228	
FT	DOMAIN 229 246	
FT	DOMAIN 247 265	
FT	DOMAIN 266 274	
FT	DOMAIN 275 390	
FT	DOMAIN 391 435	
FT	DOMAIN 436 544	
FT	DOMAIN 441 538	
FT	DOMAIN 464 469	
FT	DOMAIN 470 479	
FT	DOMAIN 480 489	
FT	DOMAIN 490 499	
FT	DOMAIN 500 509	
FT	DOMAIN 510 519	
FT	DOMAIN 520 529	
FT	DOMAIN 530 539	
FT	DOMAIN 540 549	
FT	DOMAIN 550 559	
FT	DOMAIN 560 569	
FT	DOMAIN 570 579	
FT	DOMAIN 580 589	
FT	DOMAIN 590 599	
FT	DOMAIN 600 609	
FT	DOMAIN 610 619	
FT	DOMAIN 620 629	
FT	DOMAIN 630 639	
FT	DOMAIN 640 649	
FT	DOMAIN 650 659	
FT	DOMAIN 660 669	
FT	DOMAIN 670 679	
FT	DOMAIN 680 689	
FT	DOMAIN 690 699	
FT	DOMAIN 700 709	
FT	DOMAIN 710 719	
FT	DOMAIN 720 729	
FT	DOMAIN 730 739	
FT	DOMAIN 740 749	
FT	DOMAIN 750 759	
FT	DOMAIN 760 769	
FT	DOMAIN 770 779	
FT	DOMAIN 780 789	
FT	DOMAIN 790 799	
FT	DOMAIN 800 809	
FT	DOMAIN 810 819	
FT	DOMAIN 820 829	
FT	DOMAIN 830 839	
FT	DOMAIN 840 849	
FT	DOMAIN 850 859	
FT	DOMAIN 860 869	
FT	DOMAIN 870 879	
FT	DOMAIN 880 889	
FT	DOMAIN 890 899	
FT	DOMAIN 900 909	
FT	DOMAIN 910 919	
FT	DOMAIN 920 929	
FT	DOMAIN 930 939	
FT	DOMAIN 940 949	
FT	DOMAIN 950 959	
FT	DOMAIN 960 969	
FT	DOMAIN 970 979	
FT	DOMAIN 980 989	
FT	DOMAIN 990 999	
FT	DOMAIN 1000 1009	
FT	DOMAIN 1010 1019	
FT	DOMAIN 1020 1029	
FT	DOMAIN 1030 1039	
FT	DOMAIN 1040 1049	
FT	DOMAIN 1050 1059	
FT	DOMAIN 1060 1069	
FT	DOMAIN 1070 1079	
FT	DOMAIN 1080 1089	
FT	DOMAIN 1090 1099	
FT	DOMAIN 1100 1109	
FT	DOMAIN 1110 1119	
FT	DOMAIN 1120 1129	
FT	DOMAIN 1130 1139	
FT	DOMAIN 1140 1149	
FT	DOMAIN 1150 1159	
FT	DOMAIN 1160 1169	
FT	DOMAIN 1170 1179	
FT	DOMAIN 1180 1189	
FT	DOMAIN 1190 1199	
FT	DOMAIN 1200 1209	
FT	DOMAIN 1210 1219	
FT	DOMAIN 1220 1229	
FT	DOMAIN 1230 1239	
FT	DOMAIN 1240 1249	
FT	DOMAIN 1250 1259	
FT	DOMAIN 1260 1269	
FT	DOMAIN 1270 1279	
FT	DOMAIN 1280 1289	
FT	DOMAIN 1290 1299	
FT	DOMAIN 1300 1309	
FT	DOMAIN 1310 1319	
FT	DOMAIN 1320 1329	
FT	DOMAIN 1330 1339	
FT	DOMAIN 1340 1349	
FT	DOMAIN 1350 1359	
FT	DOMAIN 1360 1369	
FT	DOMAIN 1370 1379	
FT	DOMAIN 1380 1389	
FT	DOMAIN 1390 1399	
FT	DOMAIN 1400 1409	
FT	DOMAIN 1410 1419	
FT	DOMAIN 1420 1429	
FT	DOMAIN 1430 1439	
FT	DOMAIN 1440 1449	
FT	DOMAIN 1450 1459	
FT	DOMAIN 1460 1469	
FT	DOMAIN 1470 1479	
FT	DOMAIN 1480 1489	
FT	DOMAIN 1490 1499	
FT	DOMAIN 1500 1509	
FT	DOMAIN 1510 1519	
FT	DOMAIN 1520 1529	
FT	DOMAIN 1530 1539	
FT	DOMAIN 1540 1549	
FT	DOMAIN 1550 1559	
FT	DOMAIN 1560 1569	
FT	DOMAIN 1570 1579	
FT	DOMAIN 1580 1589	
FT	DOMAIN 1590 1599	
FT	DOMAIN 1600 1609	
FT	DOMAIN 1610 1619	
FT	DOMAIN 1620 1629	
FT	DOMAIN 1630 1639	
FT	DOMAIN 1640 1649	
FT	DOMAIN 1650 1659	
FT	DOMAIN 1660 1669	
FT	DOMAIN 1670 1679	
FT	DOMAIN 1680 1689	
FT	DOMAIN 1690 1699	
FT	DOMAIN 1700 1709	
FT	DOMAIN 1710 1719	
FT	DOMAIN 1720 1729	
FT	DOMAIN 1730 1739	
FT	DOMAIN 1740 1749	
FT	DOMAIN 1750 1759	
FT	DOMAIN 1760 1769	
FT	DOMAIN 1770 1779	
FT	DOMAIN 1780 1789	
FT	DOMAIN 1790 1799	
FT	DOMAIN 1800 1809	
FT	DOMAIN 1810 1819	
FT	DOMAIN 1820 1829	
FT	DOMAIN 1830 1839	
FT	DOMAIN 1840 1849	
FT	DOMAIN 1850 1859	
FT	DOMAIN 1860 1869	
FT	DOMAIN 1870 1879	
FT	DOMAIN 1880 1889	
FT	DOMAIN 1890 1899	
FT	DOMAIN 1900 1909	
FT	DOMAIN 1910 1919	
FT	DOMAIN 1920 1929	
FT	DOMAIN 1930 1939	
FT	DOMAIN 1940 1949	
FT	DOMAIN 1950 1959	
FT	DOMAIN 1960 1969	
FT	DOMAIN 1970 1979	
FT	DOMAIN 1980 1989	
FT	DOMAIN 1990 1999	
FT	DOMAIN 2000 2009	
FT	DOMAIN 2010 2019	
FT	DOMAIN 2020 2029	
FT	DOMAIN 2030 2039	
FT	DOMAIN 2040 2049	
FT	DOMAIN 2050 2059	
FT	DOMAIN 2060 2069	
FT	DOMAIN 2070 2079	
FT	DOMAIN 2080 2089	
FT	DOMAIN 2090 2099	
FT	DOMAIN 2100 2109	
FT	DOMAIN 2110 2119	
FT	DOMAIN 2120 2129	
FT	DOMAIN 2130 2139	
FT	DOMAIN 2140 2149	
FT	DOMAIN 2150 2159	
FT	DOMAIN 2160 2169	
FT	DOMAIN 2170 2179	
FT	DOMAIN 2180 2189	
FT	DOMAIN 2190 2199	
FT	DOMAIN 2200 2209	
FT	DOMAIN 2210 2219	
FT	DOMAIN 2220 2229	
FT	DOMAIN 2230 2239	
FT	DOMAIN 2240 2249	
FT	DOMAIN 2250 2259	
FT	DOMAIN 2260 2269	
FT	DOMAIN 2270 2279	
FT	DOMAIN 2280 2289	
FT	DOMAIN 2290 2299	
FT	DOMAIN 2300 2309	
FT	DOMAIN 2310 2319	
FT	DOMAIN 2320 2329	
FT	DOMAIN 2330 2339	
FT	DOMAIN 2340 2349	
FT	DOMAIN 2350 2359	
FT	DOMAIN 2360 2369	
FT	DOMAIN 2370 2379	
FT	DOMAIN 2380 2389	
FT	DOMAIN 2390 2399	
FT	DOMAIN 2400 2409	
FT	DOMAIN 2410 2419	
FT	DOMAIN 2420 2429	
FT	DOMAIN 2430 2439	
FT	DOMAIN 2440 2449	
FT	DOMAIN 2450 2459	
FT	DOMAIN 2460 2469	
FT	DOMAIN 2470 2479	
FT	DOMAIN 2480 2489	
FT	DOMAIN 2490 2499	
FT	DOMAIN 2500 2509	
FT	DOMAIN 2510 2519	
FT	DOMAIN 2520 2529	
FT	DOMAIN 2530 2539	
FT	DOMAIN 2540 2549	
FT	DOMAIN 2550 2559	
FT	DOMAIN 2560 2569	
FT	DOMAIN 2570 2579	
FT	DOMAIN 2580 2589	
FT	DOMAIN 2590 2599	
FT	DOMAIN 2600 2609	
FT	DOMAIN 2610 2619	
FT	DOMAIN 2620 2629	
FT	DOMAIN 2630 2639	
FT	DOMAIN 2640 2649	
FT	DOMAIN 2650 2659	
FT	DOMAIN 2660 2669	
FT	DOMAIN 2670 2679	
FT	DOMAIN 2680 2689	
FT	DOMAIN 2690 2699	
FT	DOMAIN 2700 2709	
FT	DOMAIN 2710 2719	
FT	DOMAIN 2720 2729	
FT	DOMAIN 2730 2739	
FT	DOMAIN 2740 2749	
FT	DOMAIN 2750 2759	
FT	DOMAIN 2760 2769	
FT	DOMAIN 2770 2779	
FT	DOMAIN 2780 2789	
FT	DOMAIN 2790 2799	
FT	DOMAIN 2800 2809	
FT	DOMAIN 2810 2819	
FT	DOMAIN 2820 2829	
FT	DOMAIN 2830 2839	
FT	DOMAIN 2840 2849	
FT	DOMAIN 2850 2859	
FT	DOMAIN 2860 2869	
FT	DOMAIN 2870 2879	
FT	DOMAIN 2880 2889	
FT	DOMAIN 2890 2899	
FT	DOMAIN 2900 2909	
FT	DOMAIN 2910 2919	
FT	DOMAIN 2920 2929	
FT	DOMAIN 2930 2939	
FT	DOMAIN 2940 2949	
FT	DOMAIN 2950 2959	
FT	DOMAIN 2960 2969	
FT	DOMAIN 2970 2979	
FT	DOMAIN 2980 2989	
FT	DOMAIN 2990 2999	
FT	DOMAIN 3000 3009	
FT	DOMAIN 3010 3019	
FT	DOMAIN 3020 3029	
FT	DOMAIN 3030 3039	
FT	DOMAIN 3040 3049	
FT	DOMAIN 3050 3059	
FT	DOMAIN 3060 3069	
FT	DOMAIN 3070 3079	
FT	DOMAIN 3080 3089	
FT	DOMAIN 3090 3099	
FT	DOMAIN 3100 3109	
FT	DOMAIN 3110 3119	
FT	DOMAIN 3120 3129	
FT	DOMAIN 3130 3139	
FT	DOMAIN 3140 3149	
FT	DOMAIN 3150 3159	
FT	DOMAIN 3160 3169	
FT	DOMAIN 3170 3179	
FT	DOMAIN 3180 3189	
FT	DOMAIN 3190 3199	
FT	DOMAIN 3200 3209	
FT	DOMAIN 3210 3219	
FT	DOMAIN 3220 3229	
FT	DOMAIN 3230 3239	
FT	DOMAIN 3240 3249	
FT	DOMAIN 3250 3259	
FT	DOMAIN 3260 3269	
FT	DOMAIN 3270 3279	
FT	DOMAIN 3280 3289	
FT	DOMAIN 3290 3299	
FT	DOMAIN 3300 3309	
FT	DOMAIN 3310 3319	
FT	DOMAIN 3320 3329	
FT	DOMAIN 3330 3339	
FT	DOMAIN 3340 3349	
FT	DOMAIN 3350 3359	
FT	DOMAIN 3360 3369	
FT	DOMAIN 3370 3379	
FT	DOMAIN 3380 3389	
FT	DOMAIN 3390 3399	
FT	DOMAIN 3400 3409	
FT	DOMAIN 3410 3419	
FT	DOMAIN 3420 3429	
FT	DOMAIN 3430 3439	
FT	DOMAIN 3440 3449	
FT	DOMAIN 3450 3459	
FT	DOMAIN 3460 3469	
FT	DOMAIN 3470 3479	
FT	DOMAIN 3480 3489	
FT	DOMAIN 3490 3499	
FT	DOMAIN 3500 3509	
FT	DOMAIN 3510 3519	
FT	DOMAIN 3520 3529	
FT	DOMAIN 3530 3539	
FT	DOMAIN 3540 3549	
FT	DOMAIN 3550 3559	
FT	DOMAIN 3560 3569	
FT	DOMAIN 3570 3579	
FT	DOMAIN 3580 3589	
FT	DOMAIN 3590 3599	
FT	DOMAIN 3600 3609	
FT	DOMAIN 3610 3619	
FT	DOMAIN 3620 3629	
FT	DOMAIN 3630 3639	
FT	DOMAIN 3640 3649	
FT	DOMAIN 3650 3659	
FT	DOMAIN 3660 3669	
FT	DOMAIN 3670 3679	
FT	DOMAIN 3680 3689	
FT	DOMAIN 3690 3699	
FT	DOMAIN 3700 3709	
FT	DOMAIN 3710 3719	
FT	DOMAIN 3720 3729	
FT	DOMAIN 3730 3739	
FT	DOMAIN 3740 3749	
FT	DOMAIN 3750 3759	
FT	DOMAIN 3760 3769	
FT	DOMAIN 3770 3779	
FT	DOMAIN 3780 3789	
FT	DOMAIN 3790 3799	
FT	DOMAIN 3800 3809	
FT	DOMAIN 3810 3819	
FT	DOMAIN 3820 3829	
FT	DOMAIN 3830 3839	
FT	DOMAIN 3840 3849	
FT	DOMAIN 3850 3859	
FT	DOMAIN 3860 3869	
FT	DOMAIN 3870 3879	
FT	DOMAIN 3880 3889	
FT	DOMAIN 3890 3899	
FT	DOMAIN 3900 3909	
FT	DOMAIN 3910 3919	
FT	DOMAIN 3920 3929	
FT	DOMAIN 3930 3939	
FT	DOMAIN 3940 3949	
FT	DOMAIN 3950 3959	
FT	DOMAIN 3960 3969	
FT	DOMAIN 3970 3979	
FT	DOMAIN 3980 3989	
FT	DOMAIN 3990 3999	
FT	DOMAIN 4000 4009	
FT	DOMAIN 4010 4019	
FT	DOMAIN 4020 4029	
FT	DOMAIN 4030 4039	
FT	DOMAIN 4040 4049	
FT	DOMAIN 4050 4059	
FT	DOMAIN 4060 4069	
FT	DOMAIN 4070 4079	
FT	DOMAIN 4080 4089	
FT	DOMAIN 4090 4099	
FT	DOMAIN 4100 4109	
FT	DOMAIN 4110 4119	
FT	DOMAIN 4120 4129	
FT	DOMAIN 4130 4139	
FT	DOMAIN 4140 4149	
FT	DOMAIN 4150 4159	
FT	DOMAIN 4160 4169	
FT	DOMAIN 4170 4179	
FT	DOMAIN 4180 4189	
FT	DOMAIN 4190 4199	
FT	DOMAIN 4200 4209	
FT	DOMAIN 4210 4219	
FT	DOMAIN 4220 4229	
FT	DOMAIN 4230 4239	
FT	DOMAIN 4240 4249	

```
FT DOMAIN 93 124 COIL 1A.
FT DOMAIN 125 137 LINKER 1.
FT DOMAIN 138 233 COIL 1B.
FT DOMAIN 234 252 LINKER 12.
FT DOMAIN 253 271 COIL 2A.
FT DOMAIN 272 280 LINKER 2.
FT DOMAIN 281 396 COIL 2B.
FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
FT DOMAIN 444 541 TAIL, SUBDOMAIN B (ACIDIC).
FT CARBOHYD 20 O-LINKED (GLCNAC).
FT CARBOHYD 26 /FTID-CAR_000128.
FT CARBOHYD 381 /FTID-CAR_000129.
FT SITE 391 EPIPOPE (RECOGNIZED BY IF-SPECIFIC
FT CONFLICT 197 202 GADEAA -> KARMSS (IN REF. 2).
FT CONFLICT 399 399 R -> K (IN REF. 2).
FT CONFLICT 476 476 A -> E (IN REF. 2).
FT CONFLICT 480 483 EKER -> KKDE (IN REF. 2).
SQ SEQUENCE 541 AA; 61204 MW; 0D17839AF226918A CRC64;

Query Match 14.3%; Score 174; DB 1; Length 541;
Best Local Similarity 26.4%; Pred. No. 0.00061;
Matches 65; Conservative 40; Mismatches 89; Indels 52; Gaps 10;

Qy 20 KEALDNRSENGKLEALDTDFEFLPSKINGGLTSISDLPLKLRKLELRVSGGLEVL 79
Db 315 KTLIEACR---GMNEALERQLOLE--DKONADISAMQD---TINKLENELRSTKSEM 365
Qy 80 AERCPNTHLYLSGNKTKDLSTIEPLKQLENKSLDLFNCFVNLN----- 125
Db 366 AR-----YL---KEYQDLNLVKMALDIEAIRYKLLGEETRLSTSVGTSITSGYSOS 415
Qy 126 --DYGENVFKLLQLTYLDSK-----YNDHKEAPYSDIEDHVEGLDDEE-----EGE 170
Db 416 SQVFGRSAYSGLOSSYLSARAPPAYTTSVHVEEQEVEETATKAEEKDPPEGE 475
Qy 171 HEEYDDAOVDEEGEREEEREEEDVSGGDEDEGYNDGEVDGED---EEELGEE 227
Db 476 AEEE-EKEEKEEGEREEAEEAAKDESDAKEEGEGEEDTKESSEEEKESAGEE 534
Qy 228 ERQKR 233
Db 535 QAAKK 540

RESULT 7
CNG4_BOVIN STANDARD; PRT; 1394 AA.
ID CNG4_BOVIN STANDARD; PRT; 1394 AA.
AC Q28181; Q28082; Q03861;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 240 kDa protein of rod photoreceptor CNG-channel [Contains: Glutamic
DE acid-rich protein (GARP); Cyclic nucleotide-gated cation channel 4
DE (CNG channel 4) (CNG-4) (Cyclic nucleotide-gated cation channel
DE modulatory subunit)].
DE CNGB1 OR CNGG4.
GN Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
OX [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP MEDLINE=96009859; PubMed=7546742;
RX Koerschen H.G., Illing M., Seifert R., Sesti F., Williams A.,
RA Gotzes S., Colville C., Mueller F., Dose A., Godde M., Molday L.,
RA Kaupp U.B., Molday R.S.;
RA "A 240 kDa protein represents the complete beta subunit of the cyclic
RT nucleotide-gated channel from rod photoreceptor.";
RL Neuron 15:627-636(1995).
RN [2]
```

```
RP SEQUENCE OF 454-1394 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96198098; PubMed=8626431;
RA Biel M., Zong X., Ludwig A., Sautter A., Hofmann F.;
RT "Molecular cloning and expression of the modulatory subunit of the
RL cyclic nucleotide-gated cation channel.";
RN J. Biol. Chem. 271:6349-6355(1996).
RN [3]
RP SEQUENCE OF 1-590 FROM N.A.
RC TISSUE=Retina;
RA Sugimoto Y., Yatsunami K., Tsujimoto M., Khorana H.G., Ichikawa A.;
RL Submitted (XXX-1991) to the EMBL/Genbank/DBJ databases.
CC -!- SUBUNIT: FORMS FUNCTIONAL HETEROOLIGOMERIC CHANNELS WITH CNG3.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; CNG4C (SHOWN HERE), CNG4D
CC AND CNG4E; ARE PRODUCED BY ALTERNATIVE SPLICING. CNG4D IS BY FAR
CC THE MOST FREQUENT FORM (CNG4D:CNG4C:CNG4E = 20:2:1) IN TESTIS.
CC -!- TISSUE SPECIFICITY: RETINA, TESTIS, KIDNEY, HEART, AND BRAIN.
CC -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X89626; CAA61769.1; -
DR EMBL; X94707; CAA64367.1; -
DR EMBL; M61185; AAA30536.1; -
DR InterPro: IPR000636; M+channel_nlg.
DR InterPro: IPR000995; CNGP_binding.
DR Pfam: PF00027; CNGP_binding; 1.
DR Pfam: PF00520; Ion_trans; 1.
DR SMART; SM00100; CNGP; 1.
DR PROSITE; PS00888; CNGP_BINDING_1; 1.
DR PROSITE; PS00889; CNGP_BINDING_2; 1.
DR PROSITE; PS00442; CNGP_BINDING_3; 1.
KW Ionic channel; Ion transport; CAMP-binding; Transmembrane;
KW Multigene family; Alternative splicing.
FT CHAIN 1 590
FT CHAIN 454 1394
FT DOMAIN 1 767
FT TRANSMEM 768 800
FT DOMAIN 787 800
FT TRANSMEM 801 819
FT DOMAIN 821 844
FT TRANSMEM 845 864
FT DOMAIN 865 901
FT TRANSMEM 902 924
FT DOMAIN 925 968
FT TRANSMEM 969 988
FT DOMAIN 989 1072
FT TRANSMEM 1073 1093
FT DOMAIN 1094 1394
FT NP_BIND 1081 1219
FT BINDING 1141 1141
FT BINDING 1153 1153
FT CARBOHYD 1067 1067
FT VARSPIC 515 532
FT VARSPIC 522 530
FT CONFLICT 341 341
FT CONFLICT 454 465
FT CONFLICT 482 482
FT CONFLICT 499 499
FT CONFLICT 572 590
FT CONFLICT 1283 1283
FT CONFLICT 1289 1289
FT CONFLICT 1336 1336
FT CONFLICT 1338 1338
FT GLUTAMIC ACID-RICH PROTEIN.
FT CYCLIC-NUCLEOTIDE-GATED CATION CHANNEL 4.
FT CYTOPLASMIC (POTENTIAL).
FT H1 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT H2 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT H3 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT H4 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT H5 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT H6 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT CAMP (BY SIMILARITY).
FT CAMP (POTENTIAL).
FT N-LINKED (GLCNAC... ) (POTENTIAL).
FT MISSING (IN ISOFORM CNG4E).
FT MISSING (IN ISOFORM CNG4D).
FT K -> E (IN REF. 3).
FT REEEDEEEED -> MRAGQKGR (IN REF. 2).
FT R -> Q (IN REF. 2 AND 3).
FT A -> T (IN REF. 3).
FT VPATHEPELOVEDADADS -> GSFQMSPEALQCEALK
FT R (IN REF. 3).
FT S -> A (IN REF. 2).
FT R -> A (IN REF. 2).
FT D -> E (IN REF. 2).
FT A -> AA (IN REF. 2).
```

```
SQ SEQUENCE 1394 AA; 155064 MW; EE6DA559BE3744A7 CRC64;
Query Match 13.8%; Score 167.5; DB 1; Length 1394;
Best Local Similarity 43.5%; Pred. No. 0.0042;
Matches 40; Conservative 18; Mismatches 29; Indels 5; Gaps 3;

QY 147 DHKEAPYSDIEHVGGLDDEGEHEEVEDEDAQVVEDEGEDEEG-EEEDVSGGDEE 205
DQ 359 DEEKEDEGEDEEGREKEE-EGEKEEGRKEEKEEKEEKEEKEEKEEKEEKEEKEE 417
QY 206 DEEGYNDGVEDEEELG-----EEERGOKR 234
DQ 418 DEEGREKEEGRKEEGRKEE-EGEKEEGRKEEKEEKEEKEEKEEKEEKEEKEEKEE 449

RESULT 8
NFL_MOUSE STANDARD; PRT; 542 AA.
AC P08551;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
GN NFL OR NFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=87064433; PubMed=3785173;
RA Lewis S.A., Cowan N.J.;
RT "Anomalous placement of introns in a member of the intermediate
RL filament multigene family: an evolutionary conundrum.";
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=87158637; PubMed=3103856;
RA Julien J.-P., Meyer D., Flavell D., Hurst J., Grosfeld F.;
RT "Cloning and developmental expression of the murine neurofilament
RL gene family.";
RN [3]
RP Brain Res. 387:243-250(1986).
RP SEQUENCE OF 241-542 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=85131334; PubMed=3919033;
RA Lewis S.A., Cowan N.J.;
RT "Genetics, evolution, and expression of the 68,000-mol-wt
RL neurofilament protein: isolation of a cloned cDNA probe.";
RN [4]
RP J. Cell Biol. 100:843-850(1985).
RP SEQUENCE OF 1-27 FROM N.A.
RX MEDLINE=91060592; PubMed=2246261;
RA Nakahira K., Ikenaka K., Wada K., Tamura T.A., Furulichi T.,
RA Mikoshiba K.;
RT "Structure of the 68-kDa neurofilament gene and regulation of its
RL expression.";
RN [5]
RP J. Biol. Chem. 265:19786-19791(1990).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONFIBRILLAR
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X02165; CAB51616.1; -
CC EMBL; M20480; AAA39814.1; -
CC EMBL; M13016; AAA39810.1; -
CC EMBL; M55423; AAA39812.1; -
CC PIR; A25227; QFMSL.
CC MGD; MGI:97313; Nfl.
CC InterPro: IPR001664; IF.
CC Pfam; PF00038; filament; 1.
CC PROSITE; PS00226; IF; 1.
CC Intermediate filament; Coiled coil; Neurone; Glycoprotein.
CC INIT_MET 0 0
CC DOMAIN 1 92 HEAD.
CC DOMAIN 93 396 ROD.
CC DOMAIN 397 542 TAIL.
CC DOMAIN 93 124 COIL 1A.
CC DOMAIN 125 137 LINKER 1.
CC DOMAIN 138 233 COIL 1B.
CC DOMAIN 234 252 LINKER 12.
CC DOMAIN 253 271 COIL 2A.
CC DOMAIN 272 280 LINKER 2.
CC DOMAIN 281 396 COIL 2B.
CC DOMAIN 397 443 TAIL, SUBDOMAIN A.
CC DOMAIN 444 542 TAIL, SUBDOMAIN B (ACIDIC).
CC CARBOHYD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
CC CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
CC SITE 381 391 EPTOPE (RECOGNIZED BY IF-SPECIFIC
CC MONOCLONAL ANTIBODY).
CC CONFLICT 5 5 Y -> S (IN REF. 1).
CC CONFLICT 8 8 Y -> I (IN REF. 1).
CC CONFLICT 64 64 M -> K (IN REF. 1).
CC CONFLICT 72 72 V -> L (IN REF. 2).
CC CONFLICT 98 98 D -> H (IN REF. 1).
CC CONFLICT 194 194 R -> A (IN REF. 1).
CC CONFLICT 202 202 MISSING (IN REF. 2).
CC CONFLICT 239 239 Y -> I (IN REF. 1).
CC SEQUENCE 542 AA; 61448 MW; 8EE9B8C6F0831D8C CRC64;

Query Match 13.7%; Score 166.5; DB 1; Length 542;
Best Local Similarity 26.9%; Pred. No. 0.0017;
Matches 67; Conservative 40; Mismatches 85; Indels 57; Gaps 11;

QY 20 KELADNSRNEGKLEALTEFELEFLSKINGLTSLDPLKLRKLRVSGGLEVL 79
DQ 315 KTLTEIACR---GMNEALEKQLQLE--DKONADISAMQD---TINKLENLSTKSEM 365
QY 80 AEKCPNLTHLYSGNKIKDLSTIEPLKQLENLKSILDFNCEVTNLN----- 125
DQ 366 AR-----YL--KEYODLLNVKMDLIEIAVRKLLGEETRLSTSVGTSYGSQS 415
QY 126 --DYGENVFKLLQLTYLDSC-----YWDHKEAPYSDIEDHVEGLDDEE-----EGE 170
DQ 416 SQVGRSAYSGLOSSVYLSMARSFPAYTSHVQEETVEETIATKAEAKDEPPPEGE 475
QY 171 HEEYDEDAQVVEDEGEDEEGDEEEDVSGGDEDEGYNQGVDEDEE-----L 224
DQ 476 AEE-EEKEKEEKEEKEEKEEKEEKEEKEEKEEKEEKEEKEEKEEKEEKEEKEE 532
QY 225 GEEERGOKR 233
DQ 533 GEEQVAKKK 541

RESULT 9
YCF2_OENFI
```

```

ID YCF2_OBNPI STANDARD; PRT; 721 AA.
AC P31568;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ycf2 (ORF 2280) (Fragment).
GN YCF2.
OS Oenothera picensis (Oenothera odorata).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Myrtales; Onagraceae; Oenothera.
OX NCBI_TaxID=3946;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93169690; PubMed=8435856;
RA Ninzyk R., Shoendorf T., Hachtel W.;
RT "In-frame length mutations associated with short tandem repeats are
RT located in unassigned open reading frames of Oenothera chloroplast
RT DNA.";
RL Curr. Genet. 23:265-270(1993).
CC -!- FUNCTION: NOT YET KNOWN.
CC -!- SIMILARITY: BELONGS TO THE YCF2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64616; CAA45898.1; .
DR PIR; S29795; S29795.
KW Chloroplast; Hypothetical protein.
FT NON_TER
SQ SEQUENCE 721 AA; 82900 MW; 279A3FC1AA06F9C9 CRC64;

Query Match 13.6%; Score 165; DB 1; Length 721;
Best Local Similarity 48.2%; Pred. No. 0.0029;
Matches 41; Conservative 9; Mismatches 25; Indels 10; Gaps 3;

Qy 154 SDIEDHVGGLDDEEGEHE--EYDEDAQVVEDE--EGEEDDEEGEEDVSG-----GD 203
D 1 : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 180 SPTEEEVETGTEEEVETGTEEEVETGTEEEVETGTEEEVETGTEEEVETGTEEEVETG 239
Qy 204 EDEEGYNDGEVDGEDEEELGEE 228
D 1 : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 240 EEEVETGTEDEEGTEDEEGTEEE 264

RESULT 10
VG48_HSVSA STANDARD; PRT; 797 AA.
ID VG48_HSVSA
AC Q01033;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Hypothetical gene 48 protein.
GN 48 OR EDL45.
OS Herpesvirus saimiri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10383;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333688; PubMed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Honess R.W.;
RT "Primary structure of the herpesvirus saimiri genome.";
RL J. Virol. 66:5047-5058(1992).
CC -!- SIMILARITY: TO EBV BRF2.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64346; CAA45671.1; .
DR PIR; A36811; A36811.
KW Hypothetical protein.
SQ SEQUENCE 797 AA; 88867 MW; 9E294234AD850E23 CRC64;

Query Match 13.6%; Score 165; DB 1; Length 797;
Best Local Similarity 28.8%; Pred. No. 0.0033;
Matches 72; Conservative 37; Mismatches 89; Indels 52; Gaps 13;

Qy 22 LAL--DNSRNEGKLEALTDEFELEFLSKINGLTSIDL-----PKLKRLKE 69
D 1 : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 239 LALHPNDKASYSNILKFLTSNSHREHVTKVNVKAFMQSSLYKIIKDTEKNPSPTKMLM 298
Qy 70 LRVSG-----GLEVLAE-----KCPNLTHLYLSONKIKDLSTIEPL--KOLENLKSLDLFNC 119
D 1 : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 299 ISILSRGIGMDLFCQSQSVLKAPLIDH-----KLSPVSEYEDFDEVELCISDDEVD 352
Qy 120 EVTN---LNDYGENVFKLL--QLTYLDSYNDHKKAPYSIDIEDHVEGLDDEEGEHE--- 172
D 1 : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 353 EDGNLCVLDDSESVNSVALRQVLTVDK--QANEKEYKKIIDKSDRDRDKDEYLEN 409
Qy 173 EYV--DEDAQVVEDEGE--EEEEGEEDVSGDEDEEGYNDG-----EVGGE 218
D 1 : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 410 EYVNDDEDEGEDEDEDEKDEKDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGE 469
Qy 219 EDEEELGEE 228
D 1 : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 470 DDEDEGEDE 479

RESULT 11
NFL_HUMAN STANDARD; PRT; 543 AA.
ID NFL_HUMAN
AC P07196; Q16154;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
GN NEFL OR NFL OR NF68.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87214213; PubMed=3034332;
RA Julien J.-P., Grosveld F., Yazdankhsh K., Flavell D., Meijer D.,
RA Mushynski W.;
RT "The structure of a human neurofilament gene (NF-L): a unique exon-
RT intron organization in the intermediate filament gene family.";
RL Biochim. Biophys. Acta 909:10-20(1987).
RN [2]
RP SEQUENCE OF 1-10 FROM N.A.
RX MEDLINE=94235564; PubMed=8180132;
RA Pospelov V.A., Pospelova T.V., Julien J.-P.;
RT "AP-1 and Krox-24 transcription factors activate the neurofilament
RT light gene promoter in p19 embryonal carcinoma cells.";
RL Cell Growth Differ. 5:187-196(1994).
RN [3]
RP VARIANT CMT2E PRO-332.
RX MEDLINE=20307176; PubMed=10841809;
RA Merslyanova I.V., Perepelov A.V., Polyakov A.V., Sitnikov V.F.,
RA Dadali E.L., Oparin R.B., Petrin A.N., Evgrafov O.V.;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:53:13 ; Search time 88 Seconds
(without alignments)
547.898 Million cell updates/sec

Title: US-09-591-500A-4
Perfect score: 1216
Sequence: 1 MEMGRIHSELNRAPSDVK.....VDGEDEELGEERGOKRK 234

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.rodent:*
- 12: sp.virus:*
- 13: sp.vertibrate:*
- 14: sp.unclassified:*
- 15: sp.virus:*
- 16: sp.bacteriap:*
- 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1216	100.0	234	4	O43423	O43423 homo sapien
2	889	73.1	247	11	O35381	O35381 mus musculus
3	885	72.8	247	11	P97437	P97437 mus musculus
4	693.5	56.2	251	4	O92688	O92688 homo sapien
5	667	54.9	272	11	O9EST6	O9est6 rattus norv
6	655	53.9	272	11	O9EST5	O9est5 mus musculus
7	643	52.9	200	11	O9CXT4	O9cxt4 mus musculus
8	618	50.8	268	4	O9BTT0	O9btt0 homo sapien
9	614	50.5	260	11	P97822	P97822 mus musculus
10	610	50.2	266	4	O8WWW9	O8www9 homo sapien
11	609	50.1	260	11	O9CZD2	O9czd2 mus musculus
12	490	40.3	131	4	O95626	O95626 homo sapien
13	402	33.1	261	5	O9W895	O9w895 drosophila
14	294.5	24.2	447	10	O9SCQ7	O9scq7 arabidopsis
15	243	20.0	229	5	O62220	O62220 caenorhabdi
16	230	18.9	225	5	O01615	O01615 caenorhabdi

17	176	14.5	409	10	O9LN21	O9ln21 arabidopsis
18	174	14.3	567	4	O9HD28	O9hd28 homo sapien
19	173.5	14.3	1152	5	O8T216	O8t216 dictyosteli
20	171	14.1	1036	12	O9DUM3	O9dum3 kaposi's sa
21	170.5	14.0	792	12	O9DTL7	O9ytl7 ateline her
22	170	14.0	976	12	O9DUN0	O9dun0 kaposi's sa
23	168	13.8	1089	12	O40947	O40947 kaposi's sa
24	168	13.8	1129	12	O9QR71	O9qr71 kaposi's sa
25	165	13.6	896	5	O62348	O62348 caenorhabdi
26	165	13.6	906	5	O45718	O45718 caenorhabdi
27	164	13.5	1019	3	O94267	O94267 schizosacch
28	163.5	13.4	641	6	O9SLM9	O9slm9 macaca fasc
29	163.5	13.4	934	6	O9GMD3	O9gmd3 bos taurus
30	163.5	13.4	1162	12	O98148	O98148 kaposi's sa
31	163	13.4	3658	10	O9W7K6	O9w7k6 arabidopsis
32	162	13.3	416	12	O9YPA9	O9ypa9 kaposi's sa
33	161.5	13.3	386	4	O8TCR7	O8tcr7 homo sapien
34	161	13.2	673	5	O9U0N1	O9u0n1 plasmodium
35	161	13.2	1003	12	O9ILX9	O9ilx9 kaposi's sa
36	161	13.2	1127	3	O9P571	O9p571 neurospora
37	160	13.2	623	4	O96J72	O96j72 homo sapien
38	160	13.2	798	13	O90307	O90307 carassius a
39	159.5	13.1	1300	12	O36421	O36421 alceaphine
40	157.5	13.0	538	11	O9ET15	O9et15 mus musculu
41	157	12.9	495	12	O918P0	O918p0 ovine herpe
42	157	12.9	1021	4	O15451	O15451 homo sapien
43	157	12.9	1251	4	O15450	O15450 homo sapien
44	156.5	12.9	791	13	O9DGL1	O9dgl1 fugu rubrip
45	156	12.8	81	10	O9LZR7	O9lzt7 arabidopsis

ALIGNMENTS

RESULT 1

O43423
ID O43423 PRELIMINARY; PRT; 234 AA.
AC O43423;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Pp32rl.
GN Pp32rl.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kadkol S.S., Brody J.R., Pevsner J., Bai J., Pasternack G.R.;
RT "Modulation of oncogenic potential by alternative gene use in human
prostate cancer.";
RL Nat. Med. 0:0-0(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Kochevar G.J., Brody J.R., Pasternack G.R.;
RT "The Structure of a Gene Encoding Pp32rl, a New Member of the Pp32
Family.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF008216; AAdl2746.1;
DR InterPro; IPR001611; LRR.
DR Pfam; PF00560; LRR; 1.
SQ SEQUENCE 234 AA; 26762 MW; 2515D9678EACC5C2 CRC64;

Query Match
Best Local Similarity 100.0%; Score 1216; DB 4; Length 234;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEMGRIHSELNRAPSDVKELALDNRSECKLEALTDTEFELEFLSKINGLTSIDL 60
|||||
Db 1 MEMGRIHSELNRAPSDVKELALDNRSECKLEALTDTEFELEFLSKINGLTSIDL 60
QY 61 PKLKRKLELRVSGGLEVLAEKCPNTHLYLSGNKIKDLSTIEPLKQLENKSLDLFNC 120

```

Db      61 PKLKLRLKRLVSGGLEVAELKCPNLTLYLGNKIKDLSTIEPLKLENKLSLDLNFCE 120
QY      121 VTNLNDYGENVFKLLQLTLYLDSYWDHKEAPYSYDIEDHVEGLDDEEGHEEYDADAQ 180
Db      121 VTNLNDYGENVFKLLQLTLYLDSYWDHKEAPYSYDIEDHVEGLDDEEGHEEYDADAQ 180
QY      181 VVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVGDGEDEEELGEEERQKRR 234
Db      181 VVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVGDGEDEEELGEEERQKRR 234

RESULT 2
O35381 PRELIMINARY; PRT; 247 AA.
AC O35381;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cerebellar leucine rich acidic nuclear protein.
GN ANP32 OR LAMP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP Chen T.-H., Brody J.R., Romantsev F.E., Yu J.-G., Voneiff E.,
RA Kayler A.E., Kuhajda F.P., Pasternack G.R.;
RT "Structure of pp32, an Acidic Nuclear Protein Which Inhibits Oncogene-
RT Induced Formation of Transformed Foci.";
RL Mol. Biol. Cell 0:0-0(1996).
DR EMBL; U73478; AAB39707.1; -.
DR MGD; MGI:108447; Anp32.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR Pfam; PF00560; LRR; 2.
DR SMART; SM00446; LRRcap; 1.
SQ SEQUENCE 247 AA; 28521 MW; 90FDCDE53CCC2918 CRC64;

Query Match 72.8%; Score 885; DB 11; Length 247;
Best Local Similarity 77.7%; Pred. No. 1.1e-53;
Matches 185; Conservative 15; Mismatches 32; Indels 6; Gaps 3;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGNEKLEALTDDEFEFLSKINGLTSISDL 60
Db 1 MEMDKRIYELNRNTPSDVKELVLDNCKSIEGKIEGLTDEFEFLSKINGLTSISNL 60

QY 61 PKL-KLRKLEL---RVSGGLEVAELKCPNLTLYLGNKIKDLSTIEPLKLENKLSLDL 116
Db 61 PKLKLRLKLELSENRIISGDLVLAELKCPNLTLYLGNKIKDLSTIEPLKLENKLSLDL 120

QY 117 FNCVETNLNDYGENVFKLLQLTLYLDSYWDHKEAPYSYDIEDHVEGLDDEEGHEEYD 176
Db 121 FNCVETNLNAYRENVEFKLLPQVYLDGYDRNKEAPSDSVEGYVE--DDDEDEDEEYD 178

QY 177 EDAQVVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVGDGEDEEELGEEERQKRR 234
Db 179 EYAQLVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVGDGEDEEELGEEERQKRR 236

RESULT 4
Q92688 PRELIMINARY; PRT; 251 AA.
AC Q92688; O00655; P78459; P78458;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE PHAPI2A protein (APRIL) (PHAPI2B protein) (Acidic protein rich in
DE leucines) (Similar to acidic protein rich IN LEUCINES).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP Vaesen M., Barnikol-Watanabe S., Kratzin H.D., Hilschmann N.;
RA Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RA Zhu L., Henning D., Valdez B.C.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 3-251 FROM N.A.
RC TISSUE-PANCREAS;
RA Mencinger M., Panagopoulos I., Contreras J.A., Mitelman F., Aman P.;
RL Genomics 0:0-0(0).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-MUSCLE;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
```

```

Db      61 PKLKLRLKRLVSGGLEVAELKCPNLTLYLGNKIKDLSTIEPLKLENKLSLDLNFCE 120
QY      121 VTNLNDYGENVFKLLQLTLYLDSYWDHKEAPYSYDIEDHVEGLDDEEGHEEYDADAQ 180
Db      121 VTNLNDYGENVFKLLQLTLYLDSYWDHKEAPYSYDIEDHVEGLDDEEGHEEYDADAQ 180
QY      181 VVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVGDGEDEEELGEEERQKRR 234
Db      181 VVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVGDGEDEEELGEEERQKRR 234

RESULT 2
O35381 PRELIMINARY; PRT; 247 AA.
AC O35381;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cerebellar leucine rich acidic nuclear protein.
GN ANP32 OR LAMP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP Chen T.-H., Brody J.R., Romantsev F.E., Yu J.-G., Voneiff E.,
RA Kayler A.E., Kuhajda F.P., Pasternack G.R.;
RT "Structure of pp32, an Acidic Nuclear Protein Which Inhibits Oncogene-
RT Induced Formation of Transformed Foci.";
RL Mol. Biol. Cell 0:0-0(1996).
DR EMBL; U73478; AAB39707.1; -.
DR MGD; MGI:108447; Anp32.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR Pfam; PF00560; LRR; 2.
DR SMART; SM00446; LRRcap; 1.
KW Nuclear protein.
SQ SEQUENCE 247 AA; 28537 MW; 82EEDCF7ECC2918 CRC64;

Query Match 73.1%; Score 889; DB 11; Length 247;
Best Local Similarity 77.7%; Pred. No. 6.1e-53;
Matches 185; Conservative 16; Mismatches 31; Indels 6; Gaps 3;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGNEKLEALTDDEFEFLSKINGLTSISDL 60
Db 1 MEMDKRIYELNRNTPSDVKELVLDNCKSIEGKIEGLTDEFEFLSKINGLTSISNL 60

QY 61 PKL-KLRKLEL---RVSGGLEVAELKCPNLTLYLGNKIKDLSTIEPLKLENKLSLDL 116
Db 61 PKLKLRLKLELSENRIISGDLVLAELKCPNLTLYLGNKIKDLSTIEPLKLENKLSLDL 120

QY 117 FNCVETNLNDYGENVFKLLQLTLYLDSYWDHKEAPYSYDIEDHVEGLDDEEGHEEYD 176
Db 121 FNCVETNLNAYRENVEFKLLPQVYLDGYDRNKEAPSDSVEGYVE--DDDEDEDEEYD 178

QY 177 EDAQVVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVGDGEDEEELGEEERQKRR 234
Db 179 EYAQLVEDEGEHEEHEEEDVSGGDEDEEGYNDGEVGDGEDEEELGEEERQKRR 236

RESULT 3
P97437 PRELIMINARY; PRT; 247 AA.
AC P97437;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Acidic nuclear phosphoprotein pp32.
GN ANP32.
```

Query Match 54.9%; Score 667; DB 11; Length 272;
Best Local Similarity 55.5%; Pred. No. 7.4e-38;
Matches 147; Conservative 39; Mismatches 45; Indels 34; Gaps 7;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGNEKLEALDFEPEEFELSKINGLTSIDL 60
DB 1 MDKKRHHLELRNRTPAARVQLVDLNCKANGDKIEGLTAEFVNLFSLINVLGSVSDL 60

QY 61 PKL-KLRKLEL---RVSGGLEVLAKCENLTHLYLSGNKIKDLSTIEPLKOLENLKSLDL 116
DB 61 PKLPKLKLESENIFGGLDLAELPSLTHLNSGNLKADISTLEPLKRLCLKSLDL 120

QY 117 FNCVTNLNDYGENVFLLIOLTYLDSYWDHKAPYSIDIE-----DHVEGLDD 165
DB 121 FCCEVTNRSDYRETVFRLLPQLSYLDGDRDQAPSDVEVDSEAPSDGEVDGVK 180

QY 166 EEEGE-----HEEYEDDAQVDEDEEEEEEEDEEYDSGGDEED-----EEGYNDGEV 215
DB 181 EEDEEGEDEEEDEEDG---EEDEDEDEDEDEDEVEDGEDEDEVEEFGHGDEV 237

QY 216 DGEEDDEELGEEE-----RGOKRK 234
DB 238 DEDEDEDEDEDEEEESGKGKRK 262

RESULT 6
Q9EST5 PRELIMINARY; PRT; 272 AA.

ID Q9EST5 AC Q9EST5; DT 01-WAR-2001 (TREMBLrel. 16, Created)
DT 01-WAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Proliferation related acidic leucine rich protein PAL31 (Similar to
DE acidic protein rich in leucines).
GN PAL31.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C JCL; TISSUE=BRAIN;
RA Mutai H., Toyoshima Y., Sun W., Takigawa H., Tanaka S., Shiota K.;
RT "Expression of a Novel Gene, Proliferation Related Acidic Leucine Rich
RT Protein (PAL31), in the Developing Brain."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB025582; BAB12436.1; -
DR EMBL; BC005628; AAO05628.1; -
DR EMBL; BC003489; AAH03489.1; -
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR Pfam; PF00560; LRR; 2.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00446; LRRcap; 1.
SQ SEQUENCE 272 AA; 31078 MW; BC22DC591AA6BC15 CRC64;

Query Match 53.9%; Score 655; DB 11; Length 272;
Best Local Similarity 54.7%; Pred. No. 4.8e-37;
Matches 145; Conservative 40; Mismatches 46; Indels 34; Gaps 7;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGNEKLEALDFEPEEFELSKINGLTSIDL 60
DB 1 MDKKRHHLELRNRTPAARVQLVDLNCKANGDKIEGLTAEFVNLFSLINVLGSVSDL 60

QY 61 PKL-KLRKLEL---RVSGGLEVLAKCENLTHLYLSGNKIKDLSTIEPLKOLENLKSLDL 116
DB 61 PKLPKLKLESENIFGGLDLAELPSLTHLNSGNLKADISTLEPLKRLCLKSLDL 120

QY 117 FNCVTNLNDYGENVFLLIOLTYLDSYWDHKAPYSIDIEHDHVEGLDDREEGH-EFFY 175
DB 121 FNCVTNLNDYSESVFKLLPQUTYLGDYDRDQAPSDAE--VDGVDEEEDEEGEED 178

QY 176 DEDAQQVDEDEEEE---EEGEERDVSG-----GDEDEGYNDGEVDGDEDEDELG 225
DB 179 DEE-----DEDEGEFEDEDEDEDVGEDDDDEVSEEEBFGLDEDEDEDEE-E 232

QY 226 EERGOKRK 234
DB 233 EGKGKGRK 241

RESULT 5
Q9EST6 PRELIMINARY; PRT; 272 AA.

ID Q9EST6 AC Q9EST6; DT 01-WAR-2001 (TREMBLrel. 16, Created)
DT 01-WAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Proliferation related acidic leucine rich protein PAL31.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR-IMAMICHI; TISSUE=BRAIN;
RX MEDLINE=21092721; PubMed=11162633;
RA Sun W., Hattori N., Mutai H., Toyoshima Y., Kimura H., Tanaka S.,
RA Shiota K.;
RT "PAL31, a nuclear protein required for progression to the S phase.";
RT Biochem. Biophys. Res. Commun. 280:1048-1054(2001).
DR EMBL; AB025581; BAB12435.1; -
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR Pfam; PF00560; LRR; 2.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00446; LRRcap; 1.
SQ SEQUENCE 272 AA; 31060 MW; 556AC9B169E9D996 CRC64;


```
RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL6J; TISSUE=CEREBELLUM;
RA Radrizzani M., Vila-Ortiz G.J., Caiferata E.G.A.,
RA Gonzalez Guerrero A.M., Di Tella M.C., Pivetta O.H., Carminatti H.,
RA Idoaga-Vargas V.P., Santa-Coloma T.A.;
RT "Differential Expression of CpD1 During Postnatal Development in Mouse
RE Cerebellum.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN RP
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL6; TISSUE=BRAIN;
RA Matsubae M.;
RT "Characterization of the nuclear transport a novel leucine-rich acidic
RL nuclear protein-like protein.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
[3]
RN RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U89345; ABA9462.2; -
DR EMBL; AB037685; BAB03507.1; -
DR EMBL; BC005690; AAH05690.1; -
DR MGD; MGI:1913721; Cla3.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR Pfam; PF00560; LRR; 2
DR SMART; SM00446; LRRcap; 1.
SQ SEQUENCE 260 AA; 29622 MW; 7F94E46D72A04780 CRC64;

Query Match 50.5%; Score 614; DB 11; Length 260;
Best Local Similarity 53.3%; Pred. No. 2.7e-34;
Matches 139; Conservative 35; Mismatches 47; Indels 40; Gaps 8;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGKLEALTFEFLSKINGLTSISDL 60
Db 1 MEMKKINLELRNRSPVEVTELVLDNCLVNGEIEGLDTFKEFLSMANVELSLARL 60
QY 61 PKL-KLRKLELR--VSGGLEVLAKCPNLTHLYLSGNKIKDLSIEPLKQLENKLSLDL 116
Db 61 PSLNKLRLKLELSDNIISGGLVLAECPNLTLYNLSGNKIKDLSIEPLKQLENKLSLDL 120
QY 117 FNCVETNLNDYGENVFKLLQLTLTLDSCYWDHKEAPYSDIEDHVEGLDDEEGEHEEYD 176
Db 121 FNCETNLNLDYRESIFELQIITYLDGFDQEDNEAPDSEEDDEGDEDEE-EEENDAG 179
QY 177 EDAQVVEDEGEDEEEDVSG-----GDEE-----DEEGYNDGEGDGE 219
Db 180 PPEGYEEDDEDEDEDEDEDEDEAGSELGEGEEVGLSYLMKEEIQDEEDDDDDYVEEGEE 239
QY 220 DEEELGEEERQKRR 234
Db 240 EEEEEEGGLGEKRR 254

RESULT 11
Q9CZD2 PRELIMINARY; PRT; 260 AA.
AC Q9CZD2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2810018A1Srik protein.
GN CLA3 OR 2810018A1SRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK012759; BAB28449.1; -
DR MGD; MGI:1913721; Cla3.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR Pfam; PF00560; LRR; 2.
DR SMART; SM00446; LRRcap; 1.
```

```
RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL6J; TISSUE=CEREBELLUM;
RA Radrizzani M., Vila-Ortiz G.J., Caiferata E.G.A.,
RA Gonzalez Guerrero A.M., Di Tella M.C., Pivetta O.H., Carminatti H.,
RA Idoaga-Vargas V.P., Santa-Coloma T.A.;
RT "Differential Expression of CpD1 During Postnatal Development in Mouse
RE Cerebellum.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN RP
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=C57BL6; TISSUE=BRAIN;
RA Matsubae M.;
RT "Characterization of the nuclear transport a novel leucine-rich acidic
RL nuclear protein-like protein.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
[3]
RN RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U89345; ABA9462.2; -
DR EMBL; AB037685; BAB03507.1; -
DR EMBL; BC005690; AAH05690.1; -
DR MGD; MGI:1913721; Cla3.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003603; LRRcap.
DR Pfam; PF00560; LRR; 2
DR SMART; SM00446; LRRcap; 1.
SQ SEQUENCE 260 AA; 29622 MW; 7F94E46D72A04780 CRC64;

Query Match 50.5%; Score 614; DB 11; Length 260;
Best Local Similarity 53.3%; Pred. No. 2.7e-34;
Matches 139; Conservative 35; Mismatches 47; Indels 40; Gaps 8;

QY 1 MEMGRIHSELNRAPSDVKELALDNRSGKLEALTFEFLSKINGLTSISDL 60
Db 1 MEMKKINLELRNRSPVEVTELVLDNCLVNGEIEGLDTFKEFLSMANVELSLARL 60
QY 61 PKL-KLRKLELR--VSGGLEVLAKCPNLTHLYLSGNKIKDLSIEPLKQLENKLSLDL 116
Db 61 PSLNKLRLKLELSDNIISGGLVLAECPNLTLYNLSGNKIKDLSIEPLKQLENKLSLDL 120
QY 117 FNCVETNLNDYGENVFKLLQLTLTLDSCYWDHKEAPYSDIEDHVEGLDDEEGEHEEYD 176
Db 121 FNCETNLNLDYRESIFELQIITYLDGFDQEDNEAPDSEED-----DDEGDEDEE-D 174
QY 177 EDAQVVEDEGE-----EEEGEEDVSG-----DEEGYNDG 213
Db 175 ED-----EDEAGPPEGYEEDDEDEDEAGSELGEGEEVGLSYLMKDEIQDEEDDDYVD- 229
QY 214 EVDGEDEEELGEEERQKRR 234
Db 230 --EGEEEEEEGLGEKRR 248

RESULT 10
Q8WWN9 PRELIMINARY; PRT; 266 AA.
AC Q8WWN9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Leucine-rich acidic protein-like protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mao Y., Xie Y., Jiang M.;
RT "Cloning and Characterization of a Novel Human LANP-L Gene.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY057381; AAL25814.1; -.
```


